



PCT

**ORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE**  
Bureau international

DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAÎTE DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

| (51) Classification internationale des brevets <sup>5</sup> :<br>C12N 15/75, 15/32, 15/90, 1/21, 15/56 //<br>(C12N 1/21, C12R 1:07, 1:25)   |   | A2   | (11) Numéro de publication internationale: WO 94/25612<br>(43) Date de publication internationale: 10 novembre 1994 (10.11.94) |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
|---|---|--|--|----------------------|---|--|----------------|----------------|------|-------|--------|------|-------|--------|------------|----|----|------------|----|-----|--------|-------|--------|------|--------|--------|
| (21) Numéro de la demande internationale: PCT/FR94/00525<br>(22) Date de dépôt international: 5 mai 1994 (05.05.94)   |   | (61) Etats désignés: AU, BR, CA, CN, HU, JP, PL, RO, RU, UA, US, brevet européen (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
| (30) Données relatives à la priorité:<br>93/05387 5 mai 1993 (05.05.93)   |   | FR   | Publiée<br><i>Sans rapport de recherche internationale, sera republiée dès réception de ce rapport.</i>                        |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
| (71) Déposants ( <i>pour tous les Etats désignés sauf US</i> ): INSTITUT PASTEUR [FR/FR]; 25-28, rue du Dr.-Roux, F-75724 Paris Cédex 15 (FR). INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE [FR/FR]; 147, rue de l'Université, F-75341 Paris Cédex 07 (FR).  |   |  |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
| (72) Inventeurs; et<br>(75) Inventeurs/Déposants ( <i>US seulement</i> ): LERECLUS, Didier [FR/FR]; 12, rue César-Frank, F-75015 Paris (FR). AGAISSE, Hervé [FR/FR]; 54, boulevard Pasteur, F-75015 Paris (FR).   |   |  |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
| (74) Mandataires: GUTMANN, Ernest etc.; Ernest Gutmann-Yves Plasseraud-Sainte-9, rue Chauveau-Lagarde, F-75008 Paris (FR).  |   |  |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
| <p><b>(54) Title:</b> NUCLEOTIDE SEQUENCES FOR THE CONTROL OF THE EXPRESSION OF DNA SEQUENCES IN A CELLULAR HOST</p> <p><b>(54) Titre:</b> SEQUENCES DE NUCLEOTIDES POUR LE CONTROLE DE L'EXPRESSION DE SEQUENCES D'ADN DANS UN HOTE CELLULAIRE</p> <p><b>(57) Abstract</b></p> <p>The invention relates to sequences of nucleotides of bacteria, particularly Gram<sup>+</sup> bacteria such as bacteria of the Bacillus type and more particularly sequences of nucleotides of the gene <i>CryIII</i>A for the control of the expression of DNA sequences in a cellular host. The invention relates particularly to an expression system comprising a DNA sequence susceptible of being involved in the control of the expression of a coding sequence of nucleotides. Said DNA sequence comprises a promoter, as well as a sequence of nucleotides called "downstream region", situated between the promoter and the coding sequence of the gene to be expressed, and susceptible of acting at the post-transcriptional level during the expression of the gene. Preferably, the downstream region comprises a nucleotide sequence S2 comprising an essentially complementary region at the extremity 3' of the RNA 16S of the ribosomes of Bacillus type bacteria.</p> <table border="1"> <thead> <tr> <th rowspan="2">PLASMIDS (PITT—lacZ)</th> <th colspan="2">β-D-GALACTOSIDASE SPECIFIC ACTIVITY (U/mg of protein)</th> </tr> <tr> <th>t<sub>0</sub></th> <th>t<sub>6</sub></th> </tr> </thead> <tbody> <tr> <td>7830</td> <td>3.000</td> <td>35.000</td> </tr> <tr> <td>7832</td> <td>3.000</td> <td>13.000</td> </tr> <tr> <td>7830 ΔMunI</td> <td>10</td> <td>30</td> </tr> <tr> <td>7830 ΔMunI</td> <td>25</td> <td>450</td> </tr> <tr> <td>7815/8</td> <td>3.000</td> <td>42.000</td> </tr> <tr> <td>7831</td> <td>10.000</td> <td>45.000</td> </tr> </tbody> </table> |   |  |  | PLASMIDS (PITT—lacZ) | β-D-GALACTOSIDASE SPECIFIC ACTIVITY (U/mg of protein) |  | t <sub>0</sub> | t <sub>6</sub> | 7830 | 3.000 | 35.000 | 7832 | 3.000 | 13.000 | 7830 ΔMunI | 10 | 30 | 7830 ΔMunI | 25 | 450 | 7815/8 | 3.000 | 42.000 | 7831 | 10.000 | 45.000 |
| PLASMIDS (PITT—lacZ)  | β-D-GALACTOSIDASE SPECIFIC ACTIVITY (U/mg of protein) |  |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
|   | t <sub>0</sub>  | t <sub>6</sub>   |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
| 7830  | 3.000   | 35.000   |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
| 7832  | 3.000   | 13.000   |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
| 7830 ΔMunI  | 10  | 30   |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
| 7830 ΔMunI  | 25  | 450  |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
| 7815/8  | 3.000   | 42.000   |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |
| 7831  | 10.000  | 45.000   |  |                      |   |  |                |                |      |       |        |      |       |        |            |    |    |            |    |     |        |       |        |      |        |        |

(57) Abrégé

L'invention a pour objet des séquences de nucléotides de bactéries, particulièrement de bactéries Gram<sup>+</sup> telles des bactéries de type Bacillus et plus particulièrement des séquences de nucléotides du gène *CryIIIA* pour le contrôle de l'expression de séquences d'ADN dans un hôte cellulaire. L'invention concerne notamment un système d'expression comprenant une séquence d'ADN susceptible d'intervenir dans le contrôle de l'expression d'une séquence codante de nucléotides. Cette séquence d'ADN comprend un promoteur, ainsi qu'une séquence de nucléotides dite "région aval", située entre le promoteur et la séquence codante du gène à exprimer, et susceptible de jouer un rôle au niveau post-transcriptionnel lors de l'expression du gène. De façon préférée, la région aval comprend une séquence nucléotidique S2 comportant une région essentiellement complémentaire à l'extrémité 3' de l'ARN 16S des ribosomes de bactéries de type Bacillus.

**UNIQUEMENT A TITRE D'INFORMATION**

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

|    |                           |    |  |    |                       |
|----|---------------------------|----|--|----|-----------------------|
| AT | Autriche                  | GB | Royaume-Uni                                | MR | Mauritanie            |
| AU | Australie                 | GE | Géorgie                                    | MW | Malawi                |
| BB | Barbade                   | GN | Guinée                                     | NE | Niger                 |
| BE | Belgique                  | GR | Grèce                                      | NL | Pays-Bas              |
| BF | Burkina Faso              | HU | Hongrie                                    | NO | Norvège               |
| BG | Bulgarie                  | IE | Irlande                                    | NZ | Nouvelle-Zélande      |
| BJ | Bénin                     | IT | Italie                                     | PL | Pologne               |
| BR | Brésil                    | JP | Japon                                      | PT | Portugal              |
| BY | Bélarus                   | KE | Kenya                                      | RO | Roumanie              |
| CA | Canada                    | KG | Kirghizistan                               | RU | Fédération de Russie  |
| CF | République centrafricaine | KP | République populaire démocratique de Corée | SD | Soudan                |
| CG | Congo                     | KR | République de Corée                        | SE | Suède                 |
| CH | Suisse                    | KZ | Kazakhstan                                 | SI | Slovénie              |
| CI | Côte d'Ivoire             | LI | Liechtenstein                              | SK | Slovaquie             |
| CM | Cameroun                  | LK | Sri Lanka                                  | SN | Sénégal               |
| CN | Chine                     | LU | Luxembourg                                 | TD | Tchad                 |
| CS | Tchécoslovaquie           | LV | Lettonie                                   | TG | Togo                  |
| CZ | République tchèque        | MC | Monaco                                     | TJ | Tadjikistan           |
| DE | Allemagne                 | MD | République de Moldova                      | TT | Trinité-et-Tobago     |
| DK | Danemark                  | MG | Madagascar                                 | UA | Ukraine               |
| ES | Espagne                   | ML | Mali                                       | US | Etats-Unis d'Amérique |
| FI | Finlande                  | MN | Mongolie                                   | UZ | Ouzbékistan           |
| FR | France                    |    |  | VN | Viet Nam              |
| GA | Gabon                     |    |  |    |                       |

**SEQUENCES DE NUCLEOTIDES POUR LE  
CONTROLE DE L'EXPRESSION DE SEQUENCES D'ADN DANS UN  
HOTE CELLULAIRE**

L'invention a pour objet des séquences de nucléotides de bactéries, particulièrement de bactéries Gram<sup>+</sup> telles des bactéries de type *Bacillus* et plus particulièrement des séquences de nucléotides du gène cryIIIA pour le contrôle de l'expression de séquences d'ADN dans un hôte cellulaire.

Le gène cryIIIA code pour une toxine spécifique des coléoptères et est faiblement exprimé dans *Bacillus thuringiensis* lorsqu'il est cloné dans un plasmide à faible nombre de copies.

*Bacillus thuringiensis* est une bactérie Gram-positive qui produit des quantités significatives de protéines sous forme de cristal ayant une activité毒ique vis-à-vis des larves d'insectes. On connaît deux groupes de protéines de cristal, établis selon les séquences d'acides aminés et les spécificités de toxicité :

- 1) la classe des toxines Cry (I, II, III, etc...) qui ont des similitudes de structures ;
- 2) la classe des toxines Cyt, qui n'est pas apparentée à la classe cry (Höfte, H., et al 1989. *Microbiol. Rev.* 53:242-255).

Ces toxines de *B. thuringiensis* sont intéressantes de façon générale en vue de l'élaboration de bio-pesticides, et également dans la mesure où la synthèse des protéines du cristal est connue pour être parfaitement coordonnée avec la phase de sporulation de l'organisme, rendant cet organisme intéressant pour l'étude de la régulation génétique chez les bactéries sporulantes Gram-positives.

Différents mécanismes impliqués dans la régulation de la synthèse des protéines du cristal de *B. thuringiensis* ont été décrits. Le haut niveau

d'expression de ces protéines est attribué, au moins en partie, à la stabilité de l'ARNm. Certains auteurs ont attribué la stabilité de cet ARNm à la présence en aval du gène de la toxine, d'une structure jouant un rôle de terminateur qui pourrait agir comme un rétro-régulateur positif en protégeant l'extrémité 3' de l'ARNm de la dégradation par les nucléases, augmentant ainsi la demi-vie des transcrits (Wong, H.C. et al 1986. Proc. Natl. Acad. Sci. USA 83:3233-3237).

On a également émis l'hypothèse de la présence de polypeptides impliqués dans la synthèse de protéines du cristal, polypeptides qui agiraient soit en dirigeant le repliement de la protéine sous forme d'une protéine ayant une conformation stable ou pour protéger ces protéines de la dégradation protéolytique.

Des études au microscope électronique et des études biochimiques de la sporulation de B. thuringiensis montrent que la production de la protéine de cristal est dépendante de la sporulation et se situe dans le compartiment de la cellule mère (Ribier, J. et al 1973. Ann. Inst. Pasteur 124A:311-344).

Récemment, 2 facteurs sigma, sigma 35 et sigma 28, qui dirigent spécifiquement la transcription des gènes cryIA, ont été isolés et caractérisés. Ces séquences d'acides aminés montrent une identité de 88 et 85 % avec les facteurs sigma E et sigma K de Bacillus subtilis, respectivement (Adams, L. F., 1991. J. Bactériol. 173:3846-3854). Ces facteurs sigma sont produits exclusivement dans les cellules sporulantes et sont capables de fonctionner dans le compartiment de la cellule mère, confirmant que l'expression des gènes de la protéine du cristal est contrôlée dans le temps et dans l'espace. Ainsi dans l'art antérieur, on a conclu que l'expression du gène au cours du temps est, au moins en partie, assurée par l'activation successive des facteurs sigma spécifiques de la sporulation. Jusqu'à présent trois groupes de promoteurs ont été

identifiés. Deux de ces groupes comprennent des promoteurs reconnus par des facteurs sigma déterminés, et selon l'art antérieur, les facteurs sigma liés au troisième groupe de promoteurs (dont celui du gène cryIIIA) n'ont pas été identifiés (Lereclus, D., et al 1989. American Society for Microbiology, Washington, D.C.).

Enfin le nombre de copies du plasmide portant le gène semble être un facteur important pour l'expression du gène cry dans B. thuringiensis. Dans la souche sauvage B. thuringiensis, les gènes cry sont localisés sur des plasmides de grande taille, présents à un faible nombre de copies.

Des expériences de clonage d'un fragment de 3 kb HindIII, cloné dans un plasmide à faible nombre de copies, conduisent à une production faible de toxines dans une souche acristalifère (cry) de B. thuringiensis. Au contraire, de grandes quantités de toxines sont synthétisées lorsque le gène est cloné dans des plasmides à haut nombre de copies (Arantes, O. et al 1991. Gene 108:115-119).

#### Résumé de l'invention

L'invention a pour objet des moyens permettant d'obtenir un haut niveau d'expression de la protéine codée par le gène cryIIIA et plus généralement des moyens permettant de contrôler le niveau d'expression de séquences d'ADN codant pour une protéine d'intérêt déterminé, dans des souches bactériennes, préféablement des souches Gram<sup>+</sup>, telles des souches de Bacillus, cette expression pouvant être obtenue lorsque la séquence d'ADN codante est située sur un vecteur en particulier sur un plasmide à faible nombre de copies.

De façon générale, l'invention concerne un système d'expression comprenant une séquence d'ADN, susceptible d'intervenir dans le contrôle de

l'expression d'une séquence codante de nucléotides, et obtenue en associant deux séquences de nucléotides distinctes et intervenant de façons différentes mais, de façon préférable, non dissociables, dans le contrôle de l'expression de la séquence codante. La première séquence de nucléotides présente une activité promotrice, alors que la deuxième séquence, initiée par l'activité promotrice de la première, intervient pour augmenter l'expression du gène. La séquence d'ADN de l'invention permet d'atteindre un haut niveau d'expression de la partie codante d'un gène dans une bactérie, particulièrement une bactérie de type Gram<sup>+</sup>.

La première séquence de nucléotides du système d'expression de la présente invention, identifiée dans le cadre de la présente demande comme étant le promoteur, comprend soit le promoteur de la souche hôte dans laquelle le gène d'intérêt à exprimer est introduit, soit un promoteur exogène, fonctionnel dans l'hôte utilisé. La deuxième séquence de nucléotides du système d'expression de l'invention, identifiée dans la présente demande comme étant la "région aval", désigne toute séquence, préférablement située entre le promoteur et la séquence codante d'un gène à exprimer, susceptible de jouer un rôle, particulièrement au niveau post-transcriptionnel, lors de l'expression du gène. Plus particulièrement, la région aval n'agit pas directement sur la traduction de la séquence codante à exprimer.

De façon préférée, la "région aval" comprend une séquence nucléotidique, particulièrement une séquence S2 ou analogue à S2, comportant une région essentiellement complémentaire à l'extrémité 3' de l'ARN, particulièrement l'ARN 16S, des ribosomes de bactéries, particulièrement de bactéries Gram<sup>+</sup> de type *Bacillus*.

Les nucléotides composant la séquence d'ADN selon l'invention peuvent être consécutifs ou non sur

l'enchaînement à partir duquel est définie la séquence d'ADN.

Dans le cadre de la présente demande, l'expression "séquence d'ADN susceptible d'intervenir dans le contrôle de l'expression d'une séquence codante de nucléotides" traduit la capacité de cette séquence d'ADN, d'initier ou d'empêcher l'expression de la séquence codante ou de réguler cette expression notamment au niveau de la quantité du produit exprimé.

Une séquence d'ADN selon l'invention est telle que la séquence codante de nucléotides qu'elle contrôle est placée immédiatement en aval, en phase de lecture avec elle ou est au contraire séparée de cette séquence d'ADN par un fragment de nucléotides.

L'invention concerne donc une séquence d'ADN de contrôle de l'expression d'une séquence codante d'un gène dans une cellule hôte, la séquence d'ADN est caractérisée en ce qu'elle comprend un promoteur, et une séquence nucléotidique ou région aval particulièrement située en aval dudit promoteur et en amont de ladite séquence codante. La séquence nucléotidique ou région aval comporte une région essentiellement complémentaire à l'extrémité 3' d'un ARN ribosomal bactérien. La séquence d'ADN de l'invention est susceptible d'intervenir pour augmenter l'expression de la séquence codante, placée en aval, dans un hôte cellulaire.

Les inventeurs ont identifié une séquence d'ADN du type décrit précédemment, susceptible d'intervenir dans le contrôle de l'expression de la séquence codante du gène cryIIIA, et permettant notamment d'obtenir un haut niveau de l'expression lorsque la séquence codante est placée sur un plasmide à faible nombre de copies.

L'invention concerne donc également une séquence d'ADN caractérisée par les propriétés suivantes :

- elle est comprise dans un enchaînement d'ADN d'une longueur de 1692 pb environ, délimité par les sites de restriction HindIII-PstI (fragment H<sub>2</sub>-P<sub>1</sub>), tel qu'obtenu par digestion partielle du fragment BamHI de 6 kb porté par le gène cryIIIA de Bacillus thuringiensis souche LM79 ;

- elle est susceptible d'intervenir dans le contrôle de l'expression d'une séquence codante de nucléotides placée en aval, dans un hôte cellulaire, particulièrement un hôte cellulaire bactérien du type Bacillus thuringiensis et/ou Bacillus subtilis.

Les sites de restriction dont il est question ci-dessus sont représentés à la figure 1.

Dans la suite du texte on utilisera les abréviations H<sub>n</sub> pour désigner le site HindIII ayant la position "n" par rapport au premier site HindIII du fragment BamHI. De même l'expression P<sub>n</sub> désigne le site PstI en position "n" par rapport au premier site PstI sur le fragment BamHI.

La séquence d'ADN définie ci-dessus, peut être isolée et purifiée par exemple à partir du plasmide portant le gène cryIIIA de Bacillus thuringiensis.

Le système d'expression de CryIIIA comprend une première séquence de nucléotides ou promoteur située entre les sites TaqI et PacI (positions 907 à 990) et une deuxième séquence de nucléotides ou "région aval" comprise entre les sites XmnI et TaqI (positions 1179 à 1559) tel que démontré à la figure 6. La présence de deux séquences de ce type est préférée pour obtenir un niveau optimal d'expression du gène cryIIIA ou d'un autre gène placé sous le contrôle de ce système d'expression.

Entre également dans le cadre de l'invention, un vecteur d'expression caractérisé en ce qu'il est modifié en l'un de ses sites, par une séquence d'ADN telle que décrite ci-dessus, de façon telle que ladite

séquence d'ADN intervient dans le contrôle de l'expression d'une séquence codante de nucléotides déterminée.

Un vecteur de l'invention peut être de préférence un plasmide par exemple un plasmide de type réplicatif.

Un vecteur particulièrement avantageux est le plasmide pHT7902' lacZ déposé à la CNCM (Collection Nationale de Cultures de Micro-organismes - Paris - France) le 20 Avril 1993 sous le N° I-1301.

L'invention a également pour objet un hôte cellulaire recombinant caractérisé en ce qu'il est modifié par une séquence d'ADN telle que définie précédemment, ou par un vecteur d'expression décrit ci-dessus. Un hôte cellulaire particulièrement avantageux est la souche 407-OA::Km<sup>R</sup> (pHT305P) déposé à la CNCM le 3 mai 1994 sous le n° I-1412.

#### Description détaillée de l'invention

L'invention a pour objet une séquence d'ADN susceptible d'influencer l'expression de la partie codante d'un gène dans une cellule hôte bactérienne. Plus particulièrement, l'invention concerne l'association de deux séquences de nucléotides, soit un promoteur et une région aval susceptible d'intervenir au niveau post-transcriptionnel lors de l'expression de la partie codante d'un gène.

Le système d'expression de l'invention qui, comme il sera décrit en détail plus loin, implique vraisemblablement l'hybridation d'une partie de la région aval à l'extrémité 3' de l'ARN 16S d'un ribosome bactérien, peut être utilisé pour l'expression de gènes dans un grand éventail de cellules hôtes. Cette utilisation étendue du système d'expression de l'invention est possible, étant donné l'importante homologie observée au niveau des divers ARN 16S de ribosomes bactériens. Les inventeurs ayant déterminé

les régions essentielles à son fonctionnement, le système d'expression de la présente invention peut donc être utilisé dans tout type de bactéries hôtes, les ajustements nécessaires faisant partie des connaissances de l'homme de métier.

De façon générale et sans vouloir s'y restreindre pour des raisons qui deviendront évidentes plus loin, le système d'expression de la présente invention, lorsqu'utilisé pour l'expression de gènes dans des bactéries Gram<sup>+</sup> de type Bacillus, est situé en amont de la partie codante du gène à exprimer. Plus particulièrement, la région aval est normalement située immédiatement en amont du gène alors que le promoteur est situé en amont de la région aval bien qu'on puisse envisager une autre position pour ce dernier. On peut envisager le déplacement de la région aval lorsque le système est utilisé dans une cellule hôte de type E. coli chez qui les ARNm sont dégradés en sens inverse. On peut également envisager l'utilisation d'une région aval en aval et en amont de la séquence codante, ce qui permettrait la "protection" de la région codante par un mécanisme qui sera décrit en détail plus loin.

Selon un premier mode de réalisation préféré de l'invention, la séquence d'ADN correspond à l'enchaînement HindIII-PstI ( $H_2-P_1$ ) ci-dessus décrit et comprenant deux séquences de nucléotides (un promoteur et une région aval) ayant des fonctions distinctes.

Selon un mode de réalisation particulièrement avantageux de l'invention, la séquence d'ADN correspond à l'enchaînement de nucléotides désignés par l'expression Seq N°1 et correspondant au fragment d'ADN comprenant les nucléotides 1 à 1692 de la séquence représentée à la figure 3.

Le promoteur et la région aval de la séquence d'ADN de l'invention sont décrits en détail ci-dessous.

### Séquences de nucléotides présentant une activité promotrice

De façon préférée, une séquence d'ADN de l'invention intervient au niveau du contrôle de la transcription.

Il s'agit ici des séquences de nucléotides identifiées précédemment comme étant le promoteur. De façon générale et tel que mentionné précédemment, le promoteur se situe en amont de la région aval et donc à une certaine distance de la région codante du gène. Cependant, on peut envisager le déplacement du promoteur tant qu'il reste localisé en amont de la région aval.

Quant à la nature du promoteur, il semble préférable d'utiliser un promoteur provenant de la cellule hôte utilisée pour l'expression du gène d'intérêt. Toutefois, il peut, dans certaines situations, être indiqué d'utiliser un promoteur exogène. Par exemple, des promoteurs tels les promoteurs des gènes degQ, λPL, lacZ, CryI, cryIV ou de l' $\alpha$ -amylène peuvent être utilisés.

Dans le contexte de la présente invention, des fragments particulièrement préférés et comprenant une région promotrice sont les fragments suivants, représentés à la figure 1 :

- l'enchaînement délimité par les sites de restriction TaqI-PaqI ; pour des raisons de commodité de language, on a désigné par PacI l'extrémité de ce fragment qui se trouve en réalité au nucléotide 990 de la séquence de la figure 3 alors que le site PacI se termine en position 985,

- ou tout fragment de cet enchaînement, conservant les propriétés de cet enchaînement vis-à-vis du contrôle de l'expression d'une séquence codante de nucléotides.

Plus particulièrement, toute partie d'au moins 10 nucléotides naturellement consécutifs ou non

de cet enchaînement et susceptible d'intervenir dans le contrôle de l'expression d'une séquence codante de nucléotides placée en aval dans un hôte cellulaire constitue une réalisation préférentielle de l'invention. Par exemple, à l'intérieur de la séquence mentionnée précédemment se trouvent les boîtes -35(TTGCAA) et -10(TAAGCT) du promoteur.

Selon un autre mode de caractérisation de l'invention, les séquences d'ADN "de contrôle" et comprenant le promoteur dont il est question ci-dessus sont caractérisées par leur enchaînement de nucléotides. A cet égard, l'invention a notamment pour objet les séquences d'ADN correspondant aux enchaînements suivants :

- la séquence d'ADN correspondant à l'enchaînement Seq N°3 correspondant au fragment comprenant les nucléotides 907 à 990 de la séquence représentée à la figure 3, ou un variant comprenant les nucléotides 907 à 985.

L'invention a aussi pour objet des séquences d'ADN hybridant dans des conditions non stringentes, telles qu'elles sont définies ci-dessous, avec l'un des enchaînements ci-dessus décrits. Dans ce cas, l'un des enchaînements ci-dessus en question est utilisé comme sonde.

#### **Séquences de la région aval**

Une séquence de l'invention, inclue dans la région aval, est choisie pour sa capacité à intervenir pour augmenter l'expression d'un gène, qui serait initiée par un promoteur situé en amont de cette séquence. Il s'agit vraisemblablement d'une séquence susceptible d'intervenir au niveau post-transcriptionnel lors de l'expression d'une séquence codante.

En effet, les résultats expérimentaux obtenus par les inventeurs semblent indiquer que l'effet post-

transcriptionnel de la région aval définie précédemment résulte, à tout le moins lors de l'expression du gène cryIIIA, de l'hybridation entre l'ARN ribosomal 16S de la cellule hôte et une séquence S2 de l'ARN messager de cryIIIA. Il semble que le ribosome ou une partie du ribosome se fixe sur cette région aval et protège ainsi l'ARNm d'une dégradation exonucléasique initiée en 5'. Cette fixation aurait donc pour effet d'augmenter la stabilité des messagers et ainsi d'augmenter le niveau d'expression du gène cloné.

Un des fragments particulièrement préféré dans le cadre de la réalisation de l'invention et pouvant être utilisé à titre de région aval est le fragment suivant, représenté à la figure 1 :

- l'enchaînement délimité par les sites de restriction XmnI-TaqI (positions 1179 à 1556),
- ou tout fragment de cet enchaînement conservant les propriétés de cet enchaînement vis-à-vis du contrôle de l'expression d'une séquence codante de nucléotides.

Selon un autre mode de caractérisation de l'invention, les séquences d'ADN "de contrôle" comprenant la région aval dont il est question ci-dessus sont caractérisées par leur enchaînement de nucléotides. A cet égard, l'invention a notamment pour objet les séquences d'ADN correspondant aux enchaînements suivants :

- la séquence d'ADN correspondant à l'enchaînement Seq N° 4 correspondant au fragment comprenant les nucléotides 1179 à 1559 de la séquence représentée à la figure 3,
- la séquence d'ADN correspondant à l'enchaînement Seq N° 5 correspondant au fragment comprenant les nucléotides 1179 à 1556 de la séquence représentée à la figure 3,
- la séquence d'ADN correspondant à l'enchaînement Seq N° 11 correspondant au fragment

comportant les nucléotides 1413 à 1556 de la séquence représentée à la figure 3,

- la séquence d'ADN correspondant à l'enchaînement Seq N° 8 correspondant au fragment comportant les nucléotides 1413 à 1461 de la séquence représentée à la figure 3,

- la séquence d'ADN correspondant à l'enchaînement Seq N° 9 correspondant au fragment d'ADN suivant :

5'-AGCTTGAAAGGAGGGATGCCTAAAAACGAAGAACTGCA-3'

3'-ACTTTCCCTCCCTACGGATTTTGCTTCTTG-5'

- la séquence d'ADN correspondant à l'enchaînement Seq N° 10 correspondant au fragment d'ADN suivant :

5'-CTTGAAAGGAGGGATGCCTAAAAACGAAGAAC-3'

3'-GAACTTCCTCCCTACGGATTTTGCTTCTTG-5'

L'invention a aussi pour objet des séquences d'ADN hybrideant dans des conditions non stringentes, telles qu'elles sont définies ci-dessous, avec l'un des enchaînements ci-dessus décrits. Dans ce cas, l'un des enchaînements ci-dessus en question est utilisé comme sonde.

Il semble que la région aval comprenne, dans un premier temps, une région dite "essentielle", suffisamment complémentaire à l'extrémité 3' d'un ARN ribosomal 16S bactérien pour permettre la fixation du ribosome sur cette région essentielle. En aval de cette région essentielle portant le site de fixation du ribosome, se situerait une deuxième région, comprenant une structure supplémentaire susceptible d'avoir un effet positif additionnel au niveau de l'expression de la séquence codante. Il est possible que cette deuxième séquence empêche le mouvement du ribosome une fois ce dernier fixé sur la région essentielle.

Par exemple, dans le système d'expression du gène CryIIIA, il semble que la séquence de nucléotides située entre les positions 1413 et 1556 de la séquence représentée à la figure 3 comprenne la région essentielle de fixation au ribosome ainsi que la deuxième région en aval du site de fixation. Bien que la deuxième région ne soit pas absolument essentielle pour obtenir une expression accrue de la séquence codante, il semble que sa suppression réduise les rendements d'expression. En effet, des résultats expérimentaux ont démontré que la suppression de la région se situant entre les nucléotides 1462 et 1556 de la séquence représentée à la figure 3 entraîne une légère diminution de l'expression de la séquence codante.

Il semble que la longueur minimale de la séquence de nucléotides permettant une fixation adéquate du ribosome soit d'environ 10 acides nucléiques. L'invention a donc également pour objet toute partie d'au moins 10 nucléotides naturellement consécutifs ou non de l'enchaînement H<sub>2</sub>-P<sub>1</sub>, susceptible de contrôler dans un hôte cellulaire de type Bacillus, l'expression d'une séquence codante de nucléotides placés en aval de cette partie de l'enchaînement H<sub>2</sub>-P<sub>1</sub>.

Dans le cas spécifique du système d'expression du gène CryIIIA, il semblerait que la séquence de la région "essentielle" incluant le site de fixation soit la suivante :

5'-GAAAGGAGG-3'  
3'-CTTCCTCC-5'

Il est possible d'effectuer des modifications mineures au site de fixation dans la mesure où l'intensité de l'interaction entre l'extrémité 3' de l'ARN 16S du ribosome et cette région "essentielle" soit suffisamment forte pour qu'il y ait hybridation

entre le ribosome et le site de fixation. A partir de calculs d'énergie d'interaction qui peuvent être effectués par l'homme de métier, des modifications au site de fixation peuvent être envisagées si l'intensité de la liaison demeure à peu près semblable à l'intensité mesurée lors de l'utilisation de la région "essentielle" naturelle.

Dans le cas du site de fixation illustré précédemment, il est possible d'envisager certaines modifications aux quatre premiers nucléotides ainsi qu'au septième nucléotide. Il semble toutefois que les nucléotides en positions 5, 6, 8 et 9 soient importants pour assurer une intensité d'interaction appropriée lors de l'hybridation à l'ARN 16S du ribosome.

L'extrémité 3' de l'ARN 16S de ribosome bactérien étant relativement bien conservée d'une espèce bactérienne à l'autre, le système d'expression de la présente invention peut donc être utilisé dans un grand nombre d'hôtes bactériens sans avoir à effectuer de modifications substantielles.

L'invention a donc également pour objet une séquence d'ADN caractérisée par les propriétés suivantes :

- elle est contenue dans un enchaînement de nucléotides hybridant dans des conditions non stringentes avec le fragment d'ADN compris entre les nucléotides 1413 et 1559 de la séquence représentée à la figure 3 ;

- elle est susceptible d'intervenir dans le contrôle de l'expression dans une cellule hôte d'une séquence codante, notamment d'une séquence codant pour un polypeptide de Bacillus, toxique vis-à-vis des insectes ou d'une séquence codant pour un polypeptide exprimé pendant la phase stationnaire de Bacillus.

Une séquence codant pour un polypeptide de Bacillus, toxique vis-à-vis des larves d'insectes est

par exemple une séquence comprise dans le gène cryIIIB de B. thuringiensis.

Une séquence d'ADN répondant à cette définition peut être identifiée en utilisant des amorces oligonucléotidiques.

Une hybridation en conditions non stringentes entre la séquence d'ADN testée et le fragment d'ADN compris entre les nucléotides 1413 et 1559 de la séquence de la figure 3 utilisée comme sonde sera effectuée comme suit :

La sonde ADN et les séquences fixées sur filtre de nitrocellulose ou sur filtre nylon sont hybridées à 42°C pendant 18 H avec agitation, en présence de formamide (30 %), de SSC 5X de la solution de Denhardt 1X. La solution de Denhardt 1X est composée de 0,02 % de ficoll, de 0,02 % de polyvinyl pyrrolidone et de 0,02 % de sérum albumine bovine. Le SSC 1X est composé de NaCl 0,15 M et de citrate de sodium 0,015 M. Après hybridation, le filtre est successivement lavé à 42°C pendant 10 minutes dans chacune des solutions suivantes :

- Formamide (30 %), SSC 5X
- SSC 2X
- SSC 1X
- SSC 0,5X

Les conditions d'hybridation décrites précédemment sont celles qui sont utilisées pour toutes les applications de la présente invention lorsque nécessaire.

Les séquences d'ADN selon l'invention peuvent être, le cas échéant, recombinées entre elles ou associées sur le vecteur en des sites différents. En particulier, on associera avantageusement sous forme d'un enchaînement unique, le fragment TaqI-PacI au fragment XmnI-TaqI définis ci-dessus ou encore le fragment TaqI-PacI à l'enchaînement Seq N° 8. De telles séquences ont la propriété avantageuse de permettre un

haut niveau d'expression (jusqu'à 60.000 unités Miller) de la séquence codante de nucléotides, niveau d'expression qui peut être constaté avec le gène de la  $\beta$ -galactosidase.

De plus, des fragments particulièrement préférés dans le cadre de la réalisation de l'invention sont les fragments suivants représentés à la figure 8B :

- l'enchaînement délimité par les sites de restriction TaqI-TaqI,
- ou tout fragment de ces enchaînements conservant les propriétés de ces enchaînements vis-à-vis du contrôle de l'expression d'une séquence codante de nucléotides.

Selon un autre mode de caractérisation de l'invention, les séquences d'ADN dont il est question ci-dessus sont caractérisées par leur enchaînement de nucléotides. A cet égard, l'invention a notamment pour objet les séquences d'ADN correspondant aux enchaînements suivants :

- L'enchaînement Seq N° 2, correspondant au fragment comprenant les nucléotides 907 à 1559 de la séquence représentée à la figure 3,
- la séquence d'ADN correspondant à l'enchaînement Seq N° 6 correspondant au fragment comprenant les nucléotides 907 à 1353 et 1413 à 1556 de la séquence représentée à la figure 3,
- la séquence d'ADN correspondant à l'enchaînement Seq N° 7 correspondant au fragment comprenant les nucléotides 907 à 990 et 1179 à 1559 de la séquence représentée à la figure 3.

L'invention a aussi pour objet des séquences d'ADN hybridant dans des conditions non stringentes, telles qu'elles sont définies ci-dessus, avec l'un des enchaînements ci-dessus décrits. Dans ce cas, l'un des enchaînements ci-dessus en question est utilisé comme sonde.

Les séquences d'ADN de l'invention peuvent être isolées et purifiées à partir de Bacillus, notamment de B. thuringiensis ; elles peuvent être préparées également par synthèse selon les techniques connues.

Entrent également dans le cadre de l'invention, les séquences d'ARN correspondant aux séquences d'ADN ci-dessus décrites.

L'invention a également pour objet une séquence d'ADN recombinant caractérisée en ce qu'elle comprend une séquence codante déterminée, sous le contrôle d'une séquence d'ADN répondant à l'une des définitions précédentes.

La capacité des ADN de l'invention à intervenir dans le contrôle de l'expression de séquences de nucléotides, peut être vérifiée par la mise en œuvre du test suivant :

- la séquence d'ADN de l'invention dont on veut évaluer la capacité à intervenir dans le contrôle de l'expression d'une séquence codante est insérée dans un plasmide à faible nombre de copies, en amont d'une séquence codante de nucléotides.

- le plasmide ainsi préparé est utilisé pour transformer (par exemple par électroporation) une souche de Bacillus thuringiensis et par exemple une souche B. thuringiensis HD1 cry<sup>B</sup> ;

- la souche de bacillus ainsi transformée est cultivée dans des conditions permettant l'expression de la séquence codante de nucléotides ;

- on détecte le produit d'expression de cette séquence codante de nucléotides, par des techniques de mesure qualitative et/ou quantitative courantes.

Pour la réalisation de ce test, la séquence codante de nucléotides sera avantageusement la séquence codante du gène cryIIIA de Bacillus thuringiensis ou, par exemple une séquence codant pour la  $\beta$ -galactosidase.

### Hôtes cellulaires

Différents types d'hôtes cellulaires peuvent être mis en oeuvre dans le cadre de l'invention. On citera à titre d'exemple Bacillus, par exemple Bacillus thuringiensis ou Bacillus subtilis. On peut également envisager d'utiliser des cellules telles que E. coli.

Dans des hôtes cellulaires capables de sporuler, la séquence codante peut être exprimée pendant la phase végétative ou la phase stationnaire de croissance ou pendant la sporulation.

Un hôte cellulaire intéressant dans le cadre de l'invention peut également être constitué par une cellule végétale ou animale.

Si cela est nécessaire ou souhaité, selon la nature de la séquence codante de nucléotides exprimée, une séquence signal peut être également insérée dans le vecteur d'expression de l'invention afin que le produit d'expression de la séquence codante soit exposé à la surface de l'hôte cellulaire, voire exporté, à partir de cet hôte cellulaire.

De façon tout à fait intéressante, on pourra utiliser des souches asporogènes de bacillus soit naturellement, soit par mutation, et en particulier des souches de Bacillus subtilis ou Bacillus thuringiensis.

Les inventeurs ayant mis en évidence que les séquences d'ADN de l'invention permettent l'expression d'une séquence codante déterminée indépendamment de la phase de sporulation de souches de type Bacillus, un hôte asporogène peut présenter l'avantage de fournir des moyens d'expression de séquences codantes susceptibles d'entrer dans des compositons biopesticides dont les effets négatifs éventuels vis-à-vis de l'environnement seraient atténués, voire supprimés.

L'hôte asporogène choisi est particulièrement avantageux pour exprimer une séquence codante pendant sa phase stationnaire de croissance, quand la séquence

codante est sous le contrôle d'une des séquences de l'invention.

Dans le cas de souches asporogènes de bacillus obtenues par mutation, un exemple illustrant l'efficacité particulière de ce type de souche pour l'expression d'une séquence codante pendant la phase stationnaire de croissance est la construction d'une souche de *B.thuringiensis* mutée pour le gène spoOA. Une souche de *B.thuringiensis* dont le gène spoOA est inactivé et portant un gène, par exemple un gène de toxine insecticide CryI, cryII, cryIII ou cryIV ou encore un gène d'intérêt industriel, dont l'expression est mise sous le contrôle du système d'expression de CryIIIA, présente des caractéristiques avantageuses. Notamment la souche *B.thuringiensis* 407.0A::Km<sup>R</sup> (pHT305P) dont la construction est décrite en détail plus loin a au moins les avantages suivants :

a) surproduction de protéines pendant la phase stationnaire de croissance ;

b) les protéines (par exemple des biopesticides) restent incluses dans la cellule et auraient ainsi une persistance accrue dans l'environnement ; et

c) les problèmes potentiels liés à la dissémination de spores sont ainsi évités.

D'autres caractéristiques et avantages de l'invention découlent des exemples qui suivent ainsi que des figures :

Figure 1 : Carte de restriction schématique des plasmides utilisés

(A) - Carte physique du vecteur navette pHT304 Les flèches au-dessus Erm<sup>R</sup> et Ap<sup>R</sup> indiquent la direction de transcription des gènes ermC et bla respectivement. La flèche et l'expression LacZp indiquent la direction de transcription à partir du promoteur du gène LacZ. ori Bt est la région de réPLICATION du plasmide pHT1030 de B. thuringiensis. (B) - Carte de restriction

simplifiée des fragments comportant le gène cryIIIA. Le fragment A est un fragment de 6 kb BamHI de B. thuringiensis LM79 ; les fragments de restriction G, P et H ont été obtenus à partir de digestion partielle par HindIII et C a été obtenu après digestion totale avec HindIII à partir du fragment A. Ces fragments ont été clonés dans pHT304 pour donner les dérivés pHT305A, pHT305G, pHT305P, pHT305H et pHT305C respectivement. Le gène cryIIIA (cadre hachuré) et la direction de transcription sont indiqués. Les nombres sous chaque site indiquent leur ordre de la gauche vers la droite.

Figure 2 : Analyse des protéines des transformants de B. thuringiensis exprimant le gène CryIIIA. Un volume identique d'échantillons (20 µl) a été chargé dans chaque puits. Les lignes 1 à 4 et 6 à 8 de B. thuringiensis Kurstaki HD1 Cry'B portant pHT305A, pHT305G, pHT305H, pHT305P, pHT305H<sub>2</sub>-H<sub>3</sub>, pHT305C et pHT304 respectivement. La ligne 5 correspond aux marqueurs de poids moléculaire (du haut vers le bas 97, 66, 60, 43 et 30 kDa). Les flèches indiquent les composants du cristal de 73 et 67 kDa.

Figure 3 : Séquence de nucléotides de l'extrémité 5' de la région en amont du gène cryIIIA.

(A) Carte physique du fragment H<sub>2</sub>-P<sub>1</sub>, (H<sub>2</sub>-H<sub>3</sub>+H<sub>3</sub>-P<sub>1</sub>) dans l'orientation 5' vers 3'. Les positions des nucléotides des deux sites HindIII (H<sub>2</sub> et H<sub>3</sub>) qui délimitent le fragment grisé, sont indiquées. Le second segment séquencé (fragment H<sub>3</sub>-P<sub>1</sub>) était le fragment entre le troisième site HindIII et le site PstI (P<sub>1</sub>). Un site ATG d'initiation de transcription de la toxine CryIIIA est montré. La numérotation des nucléotides est rapportée par rapport au fragment séquencé et non par rapport à l'initiation de transcription.

(B) - Séquence de nucléotides du fragment H<sub>2</sub>-P<sub>1</sub>. Le codon d'initiation ATG est indiqué en caractères gras et l'extrémité du transcrit majoritaire sur gel, spécifique du gène cryIIIA, correspond au T localisé en

position 1413. Un autre transcrit commence au nucléotide 983 ; il est apparemment minoritaire sur gel. La séquence comporte au moins deux répétitions inversées. La numérotation des nucléotides débute à partir du second site HindIII et se termine au site PstI montré sur la figure 3A.

Figure 4 : Représentation des plasmides PAF1, pHT304' lacZ, pHT7901' lacZ et pHT7902' lacZ.

Figure 5 : Profil d'activité  $\beta$ -galactosidase. La croissance des cellules Bt et les conditions de réalisation des échantillons ainsi que le test sont décrits dans "Matériel et Méthodes". Le temps  $t_0$  indique la fin de la phase exponentielle et  $t_n$  est le nombre d'heures avant (-) ou après le temps zéro.

Figure 6 : Carte de restriction détaillée des plasmides pHT7902' lacZ, 7903' lacZ, 7907' lacZ, 7909' lacZ, 7930' lacZ et 7931' lacZ. Ces plasmides ont été insérés dans B. thuringiensis et l'activité  $\beta$ -galactosidase a été mesurée au temps  $t_6$  de sporulation (en unités Miller). On observe des activités respectivement de 30.000, 30.000, 3.500, 2.000, 35.000 et 60.000.

Figure 7 : Activité  $\beta$ -galactosidase dans des souches de B. subtilis Spo<sup>-</sup> et Spo<sup>+</sup> ; les cultures sont faites en milieu SP.

Figure 8 : Carte de restriction schématique des constructions utilisées pour mesurer l'activité transcriptionnelle des régions du système d'expression de cryIIIA, chez B. thuringiensis souche kurstaki HD1 Cry<sup>-B</sup>.

A - Carte physique du vecteur pHT304-18Z. Les flèches indiquent la direction de transcription des gènes ermC, bla, lacZ et du promoteur placZ; et l'orientation de la réPLICATION chez E.coli (oriEc). ori1030 indique la région de réPLICATION du plasmide pHT1030 (Lereclus et Arantes, Mol. Microbiol. 1992, 7:35-46). SD indique le site de fixation des ribosomes du gène spoVG placé

devant le gène lacZ (Perkins et Youngman, 1986, Proc. Natl. Acad. Sci. USA, 83:140-144).

B - Représentation physique et activité transcriptionnelle des différentes régions du système d'expression de cryIIIA fusionnées avec le gène lacZ. La numérotation des nucléotides est établie d'après la séquence d'ADN du fragment H2-P1 présentée Fig. 3B. Les flèches indiquent la position des extrémités 5' des transcrits telles qu'elles sont identifiées par extension d'amorce. Les traits en pointillé indiquent la localisation des fragments déletés. L'activité  $\beta$ -galactosidase des différentes constructions a été mesurée aux temps  $t_0$  et  $t_6$  de la sporulation et est indiquée en unités Miller.

Figure 9 : Détermination de l'extrémité 5' du transcrit cryIIIA/lacZ produit par la souche de B.thuringiensis portant le plasmide pHT7815/8' lacZ. L'ARN total des cellules a été extrait à  $t_3$  et soumis à une expérience d'extension d'amorce à la transcriptase inverse en utilisant comme amorce l'oligonucléotide suivant : 5'-CGTAATCTTACGTCAGTAACTTCCACAG>-3'. Cet oligonucléotide est complémentaire de la région localisée entre le site de fixation des ribosomes du gène spoVG et le codon d'initiation du gène lacZ. Le même oligonucléotide a été utilisé pour déterminer la séquence nucléotidique de la région correspondante du plasmide pHT7815/8. L'extrémité 5' est numérotée d'après la séquence d'ADN du fragment H2-P1 présentée Fig. 3B.

Fig. 10 : Carte physique schématique des constructions utilisées pour mesurer l'activité post-transcriptionnelle de la région aval du système d'expression de cryIIIA, chez B.subtilis souche 168. La numérotation des nucléotides est établie d'après la séquence d'ADN du fragment H2-P1 présentée Fig. 3B. La

flèche indique la position du démarrage de transcription localisé en position +984. L'astérisque en position 1421 indique le remplacement GGA par CCC. Les traits en pointillé indiquent la localisation des fragments d'ADN déletés. L'activité  $\beta$ -galactosidase des différentes constructions a été mesurée au temps  $t_3$  de la sporulation et est indiquée en unités Miller.

Figure 11 : Séquence de nucléotides du gène spoOA de B.thuringiensis souche 407.

A - Carte de restriction schématique du fragment d'ADN de 2,4 kb portant le gène spoOA. La flèche indique l'orientation de la transcription du gène spoOA.

B - Séquence nucléotidique de la phase ouverte de lecture comportant la séquence codante du gène spoOA. Le codon d'initiation GTG est indiqué en caractère gras. Les deux sites HincII sont soulignés. Les trois points représentent le codon stop.

#### Matériel et méthodes

##### SOUCHE BACTERIENNES ET CONDITIONS DE CROISSANCE

Escherichia coli K-12 TG1 [Δ(lac-proAB) supE thi hdsD (F' traD36 proA<sup>+</sup> proB<sup>+</sup> lacI<sup>q</sup> lacZΔDM15)] Gibson, T.J. et al 1984. thesis. University of Cambridge, Cambridge a été utilisé en tant qu'hôte pour la construction des plasmides représentés à la figure 1B et pour le bactériophage M13.

E. coli MC1061 [hsdR mcrB araD139Δ (araABC-leu) 7679 Δ lacX74 galU galK rpsL thi] (Meissner, P.S. et al 1987. Proc. Natl. Acad. Sci. USA 84:4171-4175) a été utilisé en tant qu'hôte pour la construction des plasmides représentés à la figure 7.

B. thuringiensis souche LM 79 qui contient le gène cryIIIA, a été isolé et caractérisé par Chaufaux J. et al 1991. Les colloques de l'INRA. 58:317-324.

Cette souche appartient au sérotype 8 et produit des quantités de toxines similaires à celles produites par

d'autres souches de B. thuringiensis portant le gène cryIIIA (Donovan, V.P., et al 1988. Mol. Gen. Genet. 214:365-372. - Sekar, V., et al 1987. Proc. Natl. Acad. Sci. USA 84:7036-7040).

B. thuringiensis de la sous-espèce Kurstaki HD1 Cry'B a été utilisé comme hôte pour les études de régulation du gène cryIIIA. Les souches E. coli ont été cultivées à 37°C dans un milieu Luria et transformées selon la méthode décrite par Lederberg et Cohen (1974. Bacteriol. 119:1072-1074).

La souche B. thuringiensis espèce Kurstaki HD1 Cry'B a été cultivée et transformée par électroporation selon la technique décrite par Lereclus et al, (1989. FEMS Microbiol. Lett. 60:211-218).

Les concentrations en antibiotiques pour la sélection des bactéries était de 100 µg/ml pour l'ampicilline et 25 µg/ml pour l'érythromycine.

#### Constructions de plasmides

Le fragment BamHI de 6 Kb portant le gène cryIIIA et les régions adjacentes a été isolé à partir de B. thuringiensis LM79 et a été inséré dans le site unique BamHI de pUC19 pour produire pHT791, qui a été employé comme source d'ADN pour la construction des différents plasmides utilisés ici. Le plasmide pHT305A a été obtenu par insertion du fragment BamHI de 6 Kb dans le site unique BamHI du vecteur navette pHT304 (Arantes, O et Lereclus D. 1991, Gene 108:115-119) (Figure 1 A). Des échantillons du fragment BamHI de 6 Kb ont été partiellement ou complètement digérés avec HindIII et les fragments résultants ont été clonés entre les sites BamHI et HindIII ou au site HindIII de pHT304 pour donner les dérivés pHT305G, pHT305H, pHT305P et pHT305C (Figure 1). Le plasmide pHT305HΩH<sub>2</sub>H<sub>3</sub> a été obtenu en insérant le fragment H<sub>2</sub>-H<sub>3</sub> rempli aux extrémités dans le site SmaI de pHT305H

(fragment délimité respectivement par les deuxième et troisième sites HindIII du fragment de 6 kb).

Le fragment de 4,5 Kb SmaI - KpnI du plasmide pTV32 (Perkins, J.B. et al 1986. Proc. Natl. Acad. Sci. USA 83:140-144) contenant les gènes lacZ ermC a été cloné dans pEB111 (Leonhardt, H., et al 1988. J. Gen. Microbiol. 134:605-609) pour donner le plasmide pMC11. Le plasmide pHT304'lacZ, utilisé pour construire les fusions de transcription, a été obtenu en clonant le fragment de restriction DraI-SmaI de 3,2 Kb contenant le gène lacZ dépourvu de promoteur, isolé à partir de pMC11, au site unique SmaI de pHT304. Le plasmide pHT7901'lacZ a été obtenu en clonant le fragment  $H_3-P_1$  [(HindIII-PstI) voir figure 3A] entre les sites uniques HindIII et PstI de pHT304'lacZ. Le plasmide pHT7902'lacZ a été construit en clonant le fragment  $H_2-H_3$  (figure 3A) dans le site unique HindIII de pHT7901'lacZ. L'orientation du fragment  $H_2-H_3$  a été déterminée en cartographiant les sites de restriction HpaI et BalI par rapport au site PstI. Deux sites HpaI sont localisés aux positions nucléotidiques 50 et 392 ; le site BalI est localisé à la position nucléotidique 670 (figure 3). La structure générale des plasmides recombinants portant la fusion lacZ est donnée à la figure 4.

#### Manipulations d'ADN

Les procédures standard ont été utilisées pour extraire les plasmides à partir de E. coli, pour transfecter l'ADN recombinant du phage M13 et pour purifier l'ADN simple brin (Sambrook J, et al 1989. A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory - Cold Spring Harbor, N.Y.). Les enzymes de restriction, la ligase ADN T4 et la T4 polynucléotide kinase ont été utilisées selon les recommandations du fabricant. Le fragment de Klenow de l'ADN polymérase I et des désoxyribonucléosides triphosphates ont été

utilisés pour que les extrémités du fragment H<sub>2</sub>-H<sub>3</sub> soient des extrémités franches. Les fragments de restriction d'ADN ont été purifiés sur gels d'agarose en utilisant le kit Prep A gene (Bio-Rad). Les séquences de nucléotides ont été déterminées par la méthode didéoxy de terminaison de chaîne (Sanger, F., et al 1977. Proc. Natl. Acad. Sci. Vol. 70, 1993 USA 74:5463-5467) en utilisant les phages M13mp18 et M13mp19 comme matrices, ainsi que le kit Séquenase version 2.0 (US Biochemical. Cor. Cleveland Ohio) et [ $\alpha$ -<sup>35</sup>S] dATP (15 TBq ; Amersham, United Kingdom).

#### Analyse informatique

Les séquences d'ADN ont été analysées en utilisant les programmes de l'Institut Pasteur sur un ordinateur général de données MV10000.

#### Extraction de l'ARN, Extension des amorces, Analyse Northern de l'ARN et analyse en dot blot

La souche B. thuringiensis sous espèce Kurstaki HD1 Cry<sup>+</sup>B (pHT305P) a été cultivée dans un milieu HCT (Lecadet et al 1980. J. Gen. Microbiol. 121:203-212) à 30°C en agitant. Les échantillons ont été pris à t<sub>0</sub>, t<sub>3</sub>, t<sub>6</sub> et t<sub>n</sub> (t<sub>0</sub> est défini comme étant le début de la sporulation et t<sub>n</sub> indique le nombre d'heures après le début de la sporulation). Les cellules ont été récupérées par centrifugation, resuspendues dans un milieu HCO (Lecadet, M. M., et al 1980. J. Gen. Microbiol. 121:203-212) contenant 50 mM d'azide de sodium et immédiatement congelées à - 70°C jusqu'à l'extraction d'ARN (Glatron, M.F., et al 1972. Biochimie 54:1291-1301). Pour le test d'allongement de l'amorce, un premier oligonucléotide de 39 mer (3'-CTT AGG CTT GTT AGC TTC ACT TGT ACT ATG TTA TTT TTG-5') complémentaire de la région 3'-1544 à 1583-5' du gène cryIIIA a été synthétisé et son extrémité 5' a été marquée avec du [ $\gamma$ -<sup>32</sup>P] dATP (110 TBq/mmol) par la T4

polynucléotide kinase. L'oligonucléotide de 39-mer a été purifié sur une colonne Sephadex G-25 (Pharmacia) (incorporation d'environ 70 %) et pour être utilisé comme amorce, il a été mélangé avec 50 µg d'ARN total.

Un deuxième oligonucléotide de 32 mer, complémentaire la région localisée entre les positions 1090 et 1121, a été utilisé également comme amorce et a permis la détection d'un deuxième transcrit dont le démarrage de transcription est situé en position 983. Cet oligonucléotide correspond à la séquence 5' GTTAGATAAGCATTGAGGTAGAGTCCGTCCG-3'.

L'hybridation (à 30°C), l'extension de l'amorce et l'analyse des produits ont été effectuées comme décrit par Débarbouillé, M., et al (1983. J. Bacteriol. 153:1221-1227). Les amorces de 39 mer et de 32 mer ont été utilisées pour l'allongement du fragment H<sub>3</sub>-P<sub>1</sub> cloné dans M13-mp19 et pour l'allongement du fragment H<sub>2</sub>-P<sub>1</sub> cloné dans pHT7902' lacZ, respectivement. Les produits résultant des réactions ont été placés sur des gels en parallèle avec des produits de transcription, pour déterminer les extrémités 5' des transcrits.

Une analyse Northern blot a été réalisée avec de l'ARN dénaturé fractionné par électrophorèse sur des gels d'agarose contenant du formaldehyde à 1,5 % et transféré sous vide sur des membranes Hybond-N<sup>+</sup> (Amersham) dans 20 x SSC pour 1 h (1 x SSC correspond à 150 mM NaCl plus 15 mM citrate sodium, pH 7,0). Le fragment de restriction PstI-EcoRI de 874 bp (interne au gène cryIIIA) a été marqué avec du <sup>32</sup>P avec un kit de translation de coupure (Boehringer Mannheim), puis dénaturé et utilisé comme sonde. Une préhybridation a été réalisée à 42°C pendant 4 heures dans un milieu contenant 50 % de formamide-1M NaCl-1% sodium dodecyl sulfate (SDS)-10 x Denhardt's solution-50 mM Tris-HCl (pH 7,5)-0,1 % sodium PP, ADN de sperme de saumon dénaturé ( $\geq$ 100 µg/ml) et la sonde marquée ( $10^8$ cpm/µg) a

été ajoutée à la solution de préhybridation et l'incubation a été poursuivie pendant toute la nuit. La membrane a été lavée à 65°C pendant 30 minutes 2 fois dans 2 x SSC-0,5 % SDS, une fois dans 2 x SSC-0,5 % SDS, et une fois dans 0,5 x SSC-0,5 % SDS.

Des quantités égales d'ARN de cultures synchrones de B. thuringiensis sous espèce Kurstaki HD1 CryB portant les plasmides pHT305P ou pHT305H prises à t<sub>3</sub> ont été déposées sur des membranes Hybond-C Extra (Amersham) avec un appareil manifold (Schleicher & Schuel) en utilisant le protocole dot blot décrit par Sambrook et al (Sambrook, J., et al 1989. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). La sonde et les conditions d'hybridation étaient celles décrites dans les tests Northen blot.

#### Préparation du cristal et analyse

Les cellules ont été cultivées dans un milieu HCT à 30°C avec agitation pendant 48 heures et les cristaux ont été préparés selon la méthode décrite dans la publication Lecadet, M.M., et al (1992. Appl. Environ. Microbiol. 58:840-849), à l'exception du fait que la concentration en NaCl était de 150 mM. Pour l'électrophorèse en gel de polyacrylamide-SDS (PAGE), 20 µl de chaque échantillon ont été utilisés (Lereclus, D., et al 1989. (FEMS Microbiol. Lett. 66:211-218).

#### Test de détection de la β-Galactosidase

Les souches de E. coli et B. thuringiensis contenant les fusions de transcription lacZ ont été détectées en déposant sur le milieu solide, le substrat chromogène, du 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) (40 µg/ml) et des antibiotiques appropriés. Les souches isolées ont été cultivées comme indiqué et récupérées à t<sub>-2</sub>, t<sub>-1</sub>, t<sub>0</sub>, t<sub>1,5</sub>, t<sub>3</sub>, t<sub>4,5</sub>, t<sub>6</sub>, et t<sub>7,5</sub>. Après centrifugation, les culots ont été immédiatement congelés à - 70°C (pour

prévenir l'inactivation de la  $\beta$ -galactosidase) et décongelés juste avant le traitement par ultra-sons pour détecter la  $\beta$ -galactosidase (Msadek, T., et al 1990. J. Bacteriol. 172:824-834). Les activités spécifiques présentées (exprimées en unités Miller par milligramme de protéine) correspondent aux moyennes d'au moins deux expériences indépendantes.

- RESULTATS -

L'expression du gène cryIIIA nécessite la présence d'un fragment d'ADN en amont du gène.

Arantes et Lereclus (1991. Gene 108:115-119) ont montré que le gène cryIIIA était seulement faiblement exprimé dans la souche B. thuringiensis HD1 Cry<sup>+</sup>B lorsqu'il était cloné dans un vecteur à faible nombre de copies tel que pHT304 (4 copies par équivalent chromosome).

A partir d'un fragment BamHI de 6 kb portant le gène cryIIIA et les régions adjacentes (figure 1 B), isolé à partir de la souche B. thuringiensis LM79 spécifique pour les coléoptères, on a recherché si des régions en amont du gène pouvaient être impliquées dans la régulation de l'expression de ce gène. Le fragment de 6 kb a été cloné dans le site unique BamHI du vecteur pHT304 (figure 1 A) ; des fragments obtenus après digestions partielles ou totales par HindIII du fragment BamHI de 6 kb ont été également insérés de façon indépendante dans le même plasmide pour donner les dérivés pHT305A, pHT305G, pHT305H, pHT305P et pHT305C (figure 1 B). Les cinq plasmides recombinants ont été ensuite introduits dans B. thuringiensis sous espèce Kurstaki HD1 Cry<sup>+</sup> B par électroporation, et les transformants ont été cultivés pendant deux jours à 30°C dans un milieu HCT (Lecadet, M. M., et al 1980. J. Gen. Microbiol. 121:203-212) contenant 25  $\mu$ g d'érythromycine par ml.

Des préparations de spores comportant des cristaux ont été récupérées à partir des cultures et examinées par microscopie en contraste de phases et SDS-PAGE (figure 2). Les souches recombinantes portant les vecteurs pHT305A, pHT305G et pHT305P (figure 2 lignes 1, 2 et 4 respectivement) produisaient des grandes quantités d'un cristal rhomboïde plat caractéristique des souches actives contre les larves de coléoptères. Les composants principaux de ces cristaux étaient deux peptides d'environ 73 et 67 kDa, tels que ceux préalablement décrits pour les souches de B. thuringiensis portant cryIIIA (Donovan, W.P., et al 1988. Mol. Gen. Genet. 214:365-372).

En revanche, aucune production de cristal n'a été détectée avec les souches portant pHT305H ou pHT305C (figure 2, lignes 3 et 7 respectivement). On a émis l'hypothèse que ces plasmides sont dépourvus de certains éléments présents au contraire dans les dérivés pHT305A, pHT305G et pHT305P. Cet élément supplémentaire éventuel est situé sur un fragment d'ADN de 1 kb, entre le second et le troisième site HindIII, ce fragment étant désigné par H<sub>2</sub>-H<sub>3</sub> (figure 1 B). Pour tester si son effet d'activation dépendait de sa position, le fragment H<sub>2</sub>-H<sub>3</sub> a été ligaturé à un site SmaI de pHT305H. Dans le plasmide résultant (pHT305H<sub>2</sub>H<sub>3</sub>), le fragment H<sub>2</sub>H<sub>3</sub> est localisé en aval du gène cryIIIA. La synthèse de la toxine CryIIIA de pHT305H<sub>2</sub>H<sub>3</sub> s'est avérée aussi faible qu'avec le plasmide pHT305H (figure 2 ligne 6). Cette absence d'effet pourrait être due soit à la nouvelle localisation du fragment H<sub>2</sub>-H<sub>3</sub>, cette localisation étant inappropriée ou à la désorganisation de sa structure fonctionnelle. Dans ce cas, l'élément fonctionnel démarrant à l'intérieur du fragment H<sub>2</sub>-H<sub>3</sub> serait étendu à une région au-delà du site HindIII décrit et comprendrait potentiellement la région du promoteur.

Séquençage de l'ADN et analyse

La séquence nucléotide du fragment H<sub>2</sub>-H<sub>3</sub> de 979pb du plasmide pHT791 a été déterminée (figure 3 B). De plus la séquence de 713pb s'étendant du troisième site HindIII au premier site PstI (fragment H<sub>3</sub>-P<sub>1</sub>) a été déterminée (figure 3 B). Ce second fragment porte la région en amont du promoteur, le promoteur lui-même, le site potentiel de liaison au ribosome, et les premiers 151 codons du gène cryIIIA (Sekar, V. et al, 1987 PNAS. USA 84:7036-7040). Il n'y a pas de différence entre la séquence du fragment H<sub>3</sub>-P<sub>1</sub> isolé à partir de la souche LM79 et les régions correspondantes des gènes cryIIIA isolés à partir de B. thuringiensis sous espèce tenebrionis, B. thuringiensis sous espèce san diego et la souche EG2158 (Donovan W.P. et al, Herrnstadt C. et al, Höfte H.J. et al, Sekar V. et al). Aucune séquence potentiellement codante pour une protéine autre que celle correspondant à l'extrémité 5' de cryIIIA n'a été trouvée. Cette région présente une forte proportion de bases A+T (adenine-plus-thymine) correspondant à environ 81 % entre les bases 770 et 990 et deux séquences inversées répétées. La première séquence inversée répétée est imparfaite (16 des 17 pb sont identiques) avec un centre de symétrie au nucléotide 858 et la seconde est une répétition inversée parfaite de 12 pb avec un centre de symétrie au nucléotide 1379. Les énergies libres conduisant à la formation des structures en tige boucle, calculées selon la méthode de Tinoco et al (Tinoco, J.J. et al 1973. Nature (London) New Biol. 246:40-41), étaient respectivement de - 57,7 et - 66,1 kJ/mol.

Analyse du point d'initiation et de la durée de la transcription en présence du fragment H<sub>2</sub>-H<sub>3</sub>.

Sekar et al ont cartographié le site de début de la transcription du gène cryIIIA à partir des

ARNs isolés à partir de cellules en phase précoce (stade II) et en phase intermédiaire (stades III à IV) de sporulation) en utilisant la nucléase de haricot (mung bean). Ces périodes de croissance correspondent à  $t_2$  à  $t_5$ . L'extension à partir d'amorces a été réalisée sur des ARN extraits de cellules en culture à  $t_0$ ,  $t_3$ ,  $t_6$  et  $t_9$ , pour déterminer si d'autres points d'initiation interviennent lors des phases précoces et tardives de la croissance, et pour déterminer à quel stade la transcription maximale apparaît. Un site de départ de la transcription est apparu sous forme d'un signal faiblement radioactif dans les échantillons à  $t_0$  et  $t_9$ , et ce signal s'est avéré plus intense dans les échantillons à  $t_3$  et  $t_6$ . Ce site d'initiation de transcription a été cartographié un nucléotide en amont de celui décrit par Sekar et al.

Ces résultats indiquent que le transcrit majoritaire a pour extrémité 5' le T localisé en position 1413 (Fig.3). Cependant, le T localisé en position 1413 (Fig.3) pourrait constituer l'extrémité d'un messager stable dont le vrai démarrage se situerait en amont.

Détection et analyse quantitative de l'ARN<sub>m</sub> spécifique de la toxine cryIIIA en présence du fragment H<sub>2</sub>-H<sub>3</sub>

La transcriptase inverse est satisfaisante pour l'extension à partir d'amorces pour des fragments contenant seulement 100 à 150 bases dans la mesure où cette enzyme peut s'arrêter ou s'interrompre dans des régions comportant d'importantes structures secondaires au niveau de la matrice d'ARN. Pour étudier la présence d'un site potentiel d'initiation de transcription localisé très en amont de l'extrémité 5' du gène cryIIIA, une analyse Northern blot a été réalisée. L'ARN total de la souche portant pHT305P a été récupéré à  $t_0$ ,  $t_3$ ,  $t_6$  et  $t_9$ . Les ARN ont été séparés par électrophorèse sur gels d'agarose et hybridés avec une

sonde correspondant aux fragments internes de cryIIIA marqués (fragment PstI-EcoRI de 874 pb). Dans tous les échantillons, un transcrit principal d'environ 2,5 kb a été détecté. Ceci est cohérent avec la taille du transcrit délimité par le site d'initiation de transcription décrit ci-dessus et une séquence de terminaison potentielle localisée environ 400 pb en aval du codon stop de cryIIIA, décrite par Donovan et al.

Les quantités relatives d'ARNm spécifiques de la toxine CryIIIA synthétisée par la souche portant pHT305P et par la souche portant pHT305H ont été comparées par une procédure dot blot. Des ARN isolés à partir de cultures synchrones récupérées à  $t_3$  ont été immobilisés sur une membrane de nitrocellulose et hybridés à un excès de sonde PstI-EcoRI de cryIIIA. La souche portant pHT305P contenait environ 10 à 15 fois plus d'ARNm spécifique de cryIIIA que la souche contenant pHT305H.

Production de  $\beta$ -galactosidase à partir de la fusion  $H_2-H_3::lacZ$

La synthèse relative du transcrit cryIIIA en présence et en l'absence du fragment  $H_2-H_3$  a indiqué que ce segment d'ADN régule l'expression du gène cryIIIA au niveau de la transcription plutôt qu'au niveau de la traduction. Une fusion avec le gène lacZ a été réalisée pour tester l'effet sur la transcription, produit par le fragment  $H_2-H_3$ . Le gène lacZ dépourvu de promoteur a été sous-cloné dans le site SmaI de pHT304. Le plasmide résultant pHT304' lacZ constitue un système permettant de générer des transcriptions de fusion et d'étudier leur expression dans B. thuringiensis dans des conditions approchant celles ayant lieu naturellement avec les gènes cry (plasmide à faible nombre de copies). En conséquence, le fragment  $H_3-P_1$  de 713pb a été cloné entre les sites HindIII et PstI de

pHT304' lacZ pour donner pHT7901' lacZ. Finalement, le fragment H<sub>2</sub>-H<sub>3</sub> a été cloné dans le site HindIII de pHT7901' lacZ pour donner pHT7902' lacZ, qui porte le fragment H<sub>2</sub>-H<sub>3</sub> dans son orientation d'origine par rapport au fragment H<sub>3</sub>-P<sub>1</sub> (figure 4). Les plasmides pHT304' lacZ, pHT7901' lacZ et pHT7902' lacZ ont été introduits dans B. thuringiensis sous espèce Kurstaki HD1 CryB par électroporation. Le vecteur pHT304' lacZ avait un phénotype bleu potentiellement attribué au promoteur lacZ ou à une autre région d'ADN de pUC19 agissant comme promoteur, localisée en amont des sites de clonage. La sporulation de chaque souche a été induite et des échantillons ont été pris à t<sub>2</sub> et t<sub>1</sub> (2 heures et 1 heure avant le déclenchement de la sporulation, respectivement) et de t<sub>0</sub> à t<sub>7,5</sub> à 1,5 heure d'intervalle et testés pour l'activité β-galactosidase (figure 5). L'activité β-galactosidase de la souche portant pHT304' lacZ était constante à environ 800 unités Miller de t<sub>2</sub> à t<sub>7,5</sub>. Le niveau de la production d'enzymes de la souche portant pHT7901' lacZ a monté d'environ 250 unités Miller à t<sub>2</sub> jusqu'à environ 1.200 unités Miller à t<sub>7,5</sub>, indiquant une petite mais significative augmentation de l'activité β-galactosidase pendant la sporulation (cette augmentation n'est pas apparente à cause de l'échelle utilisée sur la figure 5). Au contraire, la souche recombinante portant pHT7902' lacZ a produit beaucoup de β-galactosidase (33.000 unités Miller à t<sub>6</sub> et t<sub>7,5</sub>). Son activité β-galactosidase a augmenté d'environ 20 fois entre t<sub>0</sub> et t<sub>6</sub> (Figure 5). Le ratio des activités des souches portant pHT7901' lacZ et pHT7902' lacZ a augmenté de 8 fois pendant la phase de croissance végétative à environ 25 fois pendant la phase tardive de la sporulation.

Les résultats présentés ci-dessus et plus précisément les Figures 4 et 5, indiquent que le système d'expression de cryIIIA est fonctionnel (à bas

nombre de copies) si la région H<sub>2</sub>-H<sub>3</sub> est présente en amont de la région H<sub>3</sub>-H<sub>1</sub>. Si c'est le cas, de très hauts niveaux d'expression sont obtenus, que ce soit avec le gène cryIIIA ou avec le gène lac Z.

1) Définition précise de la région activatrice.

Des délétions du fragment H<sub>2</sub>-P<sub>1</sub> (Fig.3 A) ont montré qu'un fragment TaqI - TaqI (positions 907 à 1559, Fig.3) était suffisant pour obtenir la forte expression du gène lac Z (plasmide pHT7930'lacZ, fig.6).

De plus, une délétion interne du fragment, entre les sites PacI et XmnI (positions 990 à 1179) ne réduit pas l'expression du gène lac Z.

Cette délétion interne a fait intervenir un adaptateur entre les sites PacI et XmnI.

Les deux oligonucléotides suivants ont été synthétisés et hybridés ensemble pour construire une séquence d'ADN double brin pouvant servir d'adaptateur entre les sites PacI et XmnI :

-5' - TAAAGATATCTTGAAGCTTCACGTGTTAACACAGGCCTGCAG -3'  
-3' - TAATTCTATAGAAACTCGAAGTGCACAAATTGTCCGGACGTC -5'

L'adaptateur utilisé ici a une séquence telle que cinq nucléotides naturellement présents après le site PacI, sont reconstitués dans le plasmide pHT7931'lacZ.

En présence de cette délétion, une meilleure expression semble être obtenue en rapprochant les deux régions TaqI - PacI (positions 907 à 990) et XmnI - TaqI (positions 1179 à 1559) (plasmide pHT7931'lacZ, fig.6).

Il s'ensuit que le système d'expression de cryIIIA nécessite l'association de deux séquences d'ADN distinctes ; l'une est comprise entre les sites TaqI et PacI (positions 907 à 990), l'autre est comprise entre les sites XmnI et TaqI (positions 1179 à 1559).

Cette conclusion est renforcée par le fait qu'en l'absence de la région XmnI - TaqI (positions 1179 à 1559), la région située en amont du site XmnI n'est pas suffisante pour obtenir le haut niveau d'expression du gène lac Z (plasmide pHT7907'lacZ, fig.6). En effet, la séquence d'ADN DraI - XmnI (positions 806 à 1179) placée en amont du gène lac Z (plasmide pHT7907'lacZ) ne permet d'obtenir chez Bt (B. thuringiensis) qu'une activité  $\beta$ -galactosidase d'environ 3500 unités Miller (à comparer aux 30.000 u M obtenues avec le plasmide pHT7902'lac Z et pHT7903'lacZ).

Ce résultat confirme donc que l'association des deux séquences TaqI - PacI (positions 907 à 990) et XmnI - TaqI (positions 1179 à 1559) est nécessaire pour que le système d'expression de cryIIIA fonctionne pleinement.

L'expérience réalisée avec le fragment DraI - XmnI en amont de lacZ (plasmide pHT7907'lacZ) indique qu'une activité promotrice est incluse entre DraI et XmnI voire entre TaqI et PacI (positions 907 à 990) puisqu'une forte activité  $\beta$ -galactosidase est obtenue quand le fragment PacI-XmnI (positions 991 à 1179) est absent.

L'analyse des ARN par extension d'amorce réalisée en utilisant un oligonucléotide complémentaire de la séquence comprise entre les positions 1090 et 1121, permet en effet de détecter un démarrage de transcription dans cette région. Celui-ci est localisé en position 983 (Fig.3) ou plus probablement en position 984. Il en résulte qu'un promoteur doit être situé quelques paires de bases en amont de ce démarrage. Bien qu'il n'y ait pas d'homologie évidente avec des promoteurs connus, les boîtes -35 (TTGCAA) et -10 (TAAGCT) du promoteur se trouveraient entre les positions 945 à 980.

Un fragment d'ADN MunI - PstI (positions 952 à 1612) placé devant lacZ (plasmide pHT7909'lacZ) confère une faible activité  $\beta$ -galactosidase, comparable à celle obtenue avec le plasmide pHT7901'lacZ (Fig. 4 et 5).

Ce résultat suggère que le promoteur situé en positions 945 et 980 peut être inactivé dans une construction débutant à MunI (position 952).

On sait toutefois que la séquence minimale nécessaire à l'expression a été définie comme débutant au site TaqI (position 907).

Il résulte de ces différentes expériences, qu'une séquence d'ADN localisée entre les sites TaqI et PacI (positions 907 à 990) est requise pour obtenir une forte expression de lacZ et en conséquence une forte transcription de cryIIIA.

#### **Mesure de l'activité du promoteur amont dans le système d'expression de cryIIIA**

Pour mesurer l'activité du promoteur amont, une fusion transcriptionnelle a été construite avec le fragment d'ADN contenant ce promoteur et le gène lacZ. Pour cela le vecteur d'expression pHT304-18Z a tout d'abord été construit (Fig. 8A). Le fragment d'ADN compris entre les positions 907 et 990 a ensuite été cloné en amont du gène lacZ pour donner le plasmide pHT7832'lacZ. L'activité  $\beta$ -galactosidase est de 3.000 U/mg de protéines à  $t_0$  et de 13.000 U/mg de protéines à  $t_6$  (Fig. 8B).

Le rôle du promoteur amont dans l'activité globale du système d'expression de cryIIIA a été évalué en analysant l'effet produit par son inactivation. Le site de restriction MunI a été rempli à l'aide du fragment Klenow de l'ADN polymérase en présence de deoxynucléotides pour donner le plasmide pHT7832 $\Delta$ MunI'lacZ. Cela conduit à additionner 4

nucléotides entre les régions -35 et -10 du promoteur (CAATTAATTG versus CAATTG). L'activité  $\beta$ -galactosidase de la souche portant pHT7832 $\Delta$  MunI' lacZ était d'environ 10 U/mg de protéines à  $t_0$  et d'environ 30 U/mg de protéines à  $t_6$  (Fig. 8B). Ce résultat indique que le promoteur amont est alors inactivé. Le fragment d'ADN contenant le site MunI modifié a été introduit dans le plasmide pHT7830' lacZ pour donner le plasmide pHT7830 $\Delta$ MunI' lacZ. L'activité  $\beta$ -galactosidase de la souche portant pHT7830 $\Delta$ MunI' lacZ était d'environ 25 U/mg de protéines à  $t_0$  et d'environ 450 U/mg de protéines à  $t_6$  (Fig. 8B). Par comparaison avec la souche portant le plasmide pHT7830' lacZ, il résulte que le promoteur amont est nécessaire au fonctionnement optimal du système d'expression de cryIIIA. Le plasmide pHT7830' lacZ correspond au vecteur pHT304-18Z dans lequel est cloné le fragment TaqI contenant le système d'expression entier de cryIIIA.

**Etude du rôle de la région en aval dans le système d'expression de cryIIIA.**

Les résultats précédents confirment que le promoteur amont est nécessaire au fonctionnement optimal du système d'expression de cryIIIA ; en revanche, il n'est pas suffisant pour rendre compte de l'activité maximale du système entier. Ce dernier aspect avait été évoqué précédemment (comparer l'activité  $\beta$ -galactosidase des souches portant les plasmides pHT7832' lacZ et pHT7831' lacZ (Fig. 8B). Le plasmide pHT7831' lacZ correspond au plasmide pHT7830' lacZ dont le fragment interne PacI-XmnI est déleté. Il résulte qu'une région dite "aval" est requise pour expliquer l'activité maximale du système d'expression de cryIIIA.

Le site de démarrage de transcription du gène cryIIIA avait été précédemment localisé en position

1413, les régions -35 et -10 du promoteur putatif devaient donc être comprises entre les nucléotides 1370 et 1412 (Sekar et al., 1987, Proc. Natl. Acad. Sci. USA, 84 : 7036-7040). De façon à évaluer l'efficacité de ce promoteur putatif, nous avons construit le plasmide pHT7815/8' lacZ dans lequel le fragment d'ADN compris entre les nucléotides 1352 et 1412 a été délégué. L'activité  $\beta$ -galactosidase de la souche portant pHT7815/8' lacZ était d'environ 3.000 U/mg de protéines à  $t_0$  et d'environ 42.000 U/mg de protéines à  $t_6$  (Fig. 8B). Ce résultat indique que la région comprise entre les nucléotides 1362 et 1412 ne joue aucun rôle essentiel dans le système d'expression de cryIIIA et ne peut donc être considérée comme le promoteur du gène cryIIIA.

Une expérience d'extension d'amorce a été réalisée avec les ARN totaux extraits à  $t_3$  d'une souche de B.thuringiensis portant le plasmide pHT7815/8' lacZ. L'extrémité 5' du transcrit majoritaire est détectée, comme précédemment, en position 1413 (Fig. 9). L'ensemble de nos résultats démontrent donc que cette extrémité ne correspond pas à un démarrage de transcription mais à l'extrémité d'un transcrit stable initié en position 984 à partir d'un promoteur amont localisé dans la région d'ADN comprise entre les sites TaqI et PacI (positions 907 à 990) et défini par les régions -35 et -10 : TTGCAA et TAAGCT. L'extrémité 5' du transcrit majoritaire de cryIIIA étant invariablement en position 1413, en présence ou en absence du fragment d'ADN compris entre les positions 1362 et 1412, il résulte que cette extrémité est déterminée par la présence d'une séquence d'ADN qui se trouve en aval de la position 1413. Le rôle de cette région s'exerce donc à un niveau post-transcriptionnel. L'analyse de cette séquence aval a été faite chez B.subtilis à l'aide de fusions transcriptionnelles avec le gène lacZ. Les différentes constructions présentées

figure 10 nous ont permis de délimiter plus précisément la région aval et de mesurer son effet post-transcriptionnel:

1. Le fragment d'ADN compris entre les nucléotides 1462 et 1556 a été déléte du plasmide pHT7830' lacZ pour donner le plasmide pHT7816' lacZ. L'activité  $\beta$ -galactosidase de la souche portant pHT7816' lacZ. était d'environ 25.000 U/mg de protéines à  $t_3$ , alors que l'activité  $\beta$ -galactosidase de la souche portant pHT7830' lacZ. était d'environ 50.000 U/mg de protéines à  $t_3$  (Fig. 10).

2. Le fragment d'ADN compris entre les nucléotides 1413 et 1556 a été déléte du plasmide pHT7830' lacZ pour donner le plasmide pHT7805' lacZ. L'activité  $\beta$ -galactosidase de la souche portant pHT7805' lacZ. était d'environ 5.000 U/mg de protéines à  $t_3$  (Fig. 10).

3. Les nucléotides GGA en position 1421-1423 ont été remplacés, dans le plasmide pHT7830' lacZ, par les nucléotides CCC pour donner le plasmide pHT7830Rm' lacZ. L'activité  $\beta$ -galactosidase de la souche portant pHT7830Rm' lacZ. était d'environ 5.000 U/mg de protéines à  $t_3$  (Fig. 10).

4. Une expérience d'extension d'amorce a été réalisée avec les ARN totaux extraits à  $t_3$  d'une souche de B.thuringiensis portant le plasmide pHT7830Rm' lacZ. L'extrémité 5' du transcrit majoritaire est détectée en position 984 et aucun transcrit ayant une extrémité 5' en position 1413 n'est détecté.

Ces quatre résultats indiquent que l'effet post-transcriptionnel de la région aval est principalement dû à la séquence nucléotidique comprise entre les nucléotides 1413 et 1461. De plus, les nucléotides GGA en position 1421-1423 sont importants pour conférer l'effet post-transcriptionnel et ne pourraient être modifiés qu'en considérant le remplacement par une séquence assurant une intensité

d'interaction avec l'ARN 16S du ribosome semblable à l'intensité d'interaction mesurée pour les nucléotides GGA. Par exemple, le remplacement des nucléotides GGA par les nucléotides CCC conduit à la disparition complète de l'effet post-transcriptionnel, expliqué par une forte modification de l'intensité d'interaction entre cette portion du segment et l'ARN 16S. La région aval ainsi définie a comme particularité remarquable de comprendre une séquence nucléotidique complémentaire de l'extrémité 3' de l'ARN 16S des ribosomes.

L'effet post-transcriptionnel de cette séquence d'ADN a ensuite été évalué en utilisant un système d'expression hétérologue : la séquence d'ADN suivante (S1) :

5'-AGCTTGAAAGGAGGGATGCCTAAAAACGAAGAACTGCA-3'  
3'-ACTTTCCTCCCTACGGATTTGCTTCTTG-5'

a été synthétisée et clonée entre les sites HindIII et PstI du vecteur pHT304' lacZ pour donner le plasmide pHT304ΩRS1' lacZ. Cette séquence d'ADN est donc intercalée entre le promoteur du gène lacZ et la séquence codante du gène lacZ. L'activité  $\beta$ -galactosidase de la souche 168 de B. subtilis portant pHT304ΩRS1' lacZ était d'environ 4.000 U/mg de protéines à  $t_3$ . Il résulte que la séquence décrite ci-dessus augmente d'un facteur 4 l'expression du gène lacZ. Cette augmentation est comparable à l'augmentation due à la région comprise entre les nucléotides 1413 et 1461 qui était d'un facteur 5 (comparer l'activité  $\beta$ -galactosidase des souches de B. subtilis contenant les plasmides pHT7816' lacZ ou pHT7805' lacZ). La région d'ADN suivante est donc suffisante pour conférer l'effet post-transcriptionnel au système d'expression de cryIIIA) :

5'-CTTGAAAGGAGGGATGCCTAAAAACGAAGAAC-3'  
3'-GAACCTTCCTCCCTACGGATTTGCTTCTTG-5'

Cette séquence possède une région complémentaire de l'extrémité 3' de l'ARN 16S des ribosomes. Toutefois, d'autres éléments caractéristiques de la région aval du système d'expression de cryIIIA et qui peuvent accentuer cet effet, notamment en empêchant le mouvement du ribosome, sont vraisemblablement inclus dans la séquence nucléotidique comprise entre les positions 1462 et 1556. Leur présence semble expliquer la différence d'activité  $\beta$ -galactosidase observée entre la souche de B.subtilis contenant le plasmide pHT7830'lacZ (50.000 U/mg de protéines à t<sub>3</sub>) et la souche de B.subtilis contenant le plasmide pHT7816'lacZ (25.000 U/mg de protéines à t<sub>3</sub> ; voir Fig. 10).

Ces résultats semblent donc confirmer que l'effet post-transcriptionnel de la région aval résulte de l'hybridation entre l'ARN ribosomal 16S et la séquence S2 de l'ARN messager de cryIIIA. Il est donc probable que le ribosome ou une partie du ribosome se fixe sur cette région aval de l'ARN et le protège ainsi d'une dégradation exonucléasique initiée en 5'. Tel que mentionné précédemment, cette fixation aurait donc pour effet d'augmenter la stabilité des messagers et ainsi d'augmenter le niveau d'expression d'un gène donné. Ceci explique que l'extrémité 5' des transcrits de cryIIIA soit invariablement en position 1413, quelle que soit l'initiation de la transcription. Ce mécanisme semble aussi confirmé par l'effet positif de la séquence S1 sur un système d'expression hétérologue (plasmide pHT304'λRS1lacZ dans la souche 168 de B.subtilis).

Introduction de la fusion [système d'expression CryIIIA - LacZ] dans le chromosome de *Bacillus subtilis*.

Le vecteur pAF1 (fig4) non replicatif chez B.subtilis permet d'introduire les fusions avec le gène

reporteur LacZ dans le chromosome de B.subtilis au locus amyE. (J. Bact 1990, 172:835 - 844).

Le plasmide pHCl est obtenu par insertion du fragment HindIII-SacI (2,7 kb) du pHT7901'Lac Z entre les sites HindIII-SacI du pAF1.

Le plasmide pHC2 est obtenu par insertion du fragment HindIII-SacI (3,7 kb) du pHT7902'LacZ entre les sites HindIII et SacI du pAF1.

Les fusions sont introduites dans la souche B.s 168 trpc2 (Anagnostopoulos,C., et Spizizen,J.1961.J Bacteriol 81:741-746) (*Bacillus subtilis* 168) par transformation ; le phénotype [amy-] rend compte de l'intégration par double recombinaison.

#### Etude du système d'expression du gène cryIIIA chez B. Subtilis.

Les souches de B.subtilis obtenues après transformation et intégration des plasmides pHCl et pHC2 sont appelées respectivement :

Bs 168 [H] et Bs 168 [P]

La construction contenue dans le plasmide pHC2, c'est-à-dire portant le fragment H<sub>2</sub>-P<sub>1</sub> en amont de lac Z, a aussi été introduite dans la souche B.subtilis Δ sigE.

On obtient la souche ΔsigE en transformant une souche parentale (Spo<sup>+</sup>) avec un plasmide non réplicatif chez les bactéries gram-positives et portant un gène sigE dont une région interne est déletée. Le gène sigE a été décrit par Stragier et al 1984. Nature 312:376-378.

La souche ΔsigE est transformée avec le plasmide pHC2 et la souche résultante est ΔsigE [P]. Cette souche a été déletée pour le gène codant pour le facteur sigmaE spécifique de la sporulation. Cette souche est donc asporogène (Spo<sup>-</sup>).

De même, la souche Bs 168 [P] a été transformée avec une "cassette Km<sup>R</sup>" qui interrompt le

gène SpoOA. La souche d'où provient le gène spoOA interrompu par une "cassette Km<sup>R</sup>" est obtenue en transformant une souche parentale (Spo<sup>+</sup>) avec un plasmide non réplicatif chez les bactéries Gram-positives et portant un gène SpoOA (décrit par Ferrari, F.A et al 1985. PNAS USA 82 : 2647-2651) interrompu par un gène de résistance à la kanamycine. L'ADN chromosomique de cette souche a été utilisé pour transformer la souche Bs 168 [P].

Ainsi la souche résultante Spo<sup>+</sup> a été appelée Bs168 SpoOA [P].

Il apparaît tout d'abord que la production de  $\beta$ -galactosidase obtenue avec la souche de B.subtilis 168 [H] est très faible (<100 uM) en comparaison avec la souche 168 [P] (environ 15.000 uM). Ces résultats sont similaires à ceux obtenus chez Bt.

De plus, un résultat très surprenant a été obtenu : l'expression dans la souche BsΔsigE est identique à l'expression dans la souche Bs 168 sauvage. Ce résultat indique que le gène cryIIIA n'est pas contrôlé par un promoteur spécifique du facteur sigma E comme c'est le cas pour le gène cryIA.

Il est encore plus surprenant que l'expression dans la souche Bs SpoOA [P] soit supérieure à celle obtenue dans la souche Bs 168 [P]. Ce résultat montre que l'expression de cryIIIA est indépendante de la sporulation puisque le gène spoOA intervient au premier stade de la sporulation.

Ces résultats sont très importants pour le développement et les applications du système d'expression de cryIIIA. Ils indiquent en effet que l'on peut envisager la production des toxines insecticides ou de tout autre protéine d'intérêt commercial dans des souches Spo<sup>+</sup> de B.subtilis ou de B.thuringiensis.

Analyse de l'expression de la fusion [système d'expression CryIIIA - LacZ = pHC2] chez Bacillus subtilis en fonction du milieu de culture.

En ce qui concerne l'expression de la fusion dans les milieux 1 à 5 respectivement dont la composition est donnée ci-dessous, on peut faire les observations suivantes :

- Il y a expression (bien que faible) durant la phase végétative.

L'expression augmente à l'entrée en phase stationnaire.

- La comparaison des milieux 2 (carence en phosphate) et 5 (carence en acides aminés) montre que le système d'expression CryIIIA est activé par la carence en acides aminés.

- L'expression dans le milieu 4 montre que cette activation nécessite la présence des sels : CaCl<sub>2</sub>, MnCl<sub>2</sub>, CAF.

- L'activation est indépendante de la sporulation :

En milieu de sporulation (milieu Sp) 1 l'expression s'arrête à t<sub>2</sub>.

Dans le milieu 5 les cellules ne peuvent sporuler (le glucose inhibe la sporulation), et l'activation est maximum.

- Quand la seule source d'azote est NH<sub>4</sub><sup>+</sup>, l'activation est moins importante, l'expression reste cependant importante (milieu 3).

1/ Milieu Sp : milieu de sporulation

8 g nutrient hoth (Difco) / litre

1 mM MgSO<sub>4</sub>

13 mM KCl

10 μM MnCl<sub>2</sub>

1 μM FeSO<sub>4</sub>

1 mM CaCl<sub>2</sub>

2/ Milieu de carence en phosphate

Tampon HEPES pH7; 50 mM

1 mM MgSO<sub>4</sub>0,5 mM CaCl<sub>2</sub>10 µM MnCl<sub>2</sub>

4,4 mg/litre citrate ammonioferrique (CAF)

2 % glucose

10 mM KCl

100 mg/litre chaque acide aminé

50 mg/litre tryptophane

0,45 mM tampon phosphate pH 7

3/ Milieu minimum44 mM KH<sub>2</sub>PO<sub>4</sub>60 mM K<sub>2</sub>HPO<sub>4</sub>

2,9 mM citrate tri Na

15 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>

2 % glucose

4/ Milieu carencé en acides aminés sans CaCl<sub>2</sub>MnCl<sub>2</sub>

CAF

44 mM KH<sub>2</sub>PO<sub>4</sub>60 mM K<sub>2</sub>HPO<sub>4</sub>

2,9 mM citrate tri Na

2 % glucose

1 mM MgSO<sub>4</sub>

50 mg/litre tryptophane

0,5 hydrolysat de Caséine (HC)

5/ idem 4 en ajoutant :0,5 mM CaCl<sub>2</sub>10 µM MnCl<sub>2</sub>

4,4 mg/litre CAF

Construction d'une souche de B.thuringiensis Sp<sup>-</sup>.Clonage du gène spoOA de B.thuringiensis :L'ADN total de la souche de B.thuringiensis 407 du sérototype 1 a été purifié et digéré par l'enzyme

HindIII. Les fragments HindIII ont été ligaturés avec le vecteur pHT304 digéré par HindIII et le mélange de ligature a été utilisé pour transformer la souche 168 de B. subtilis. Les clones transformants ont été sélectionnés pour la résistance à l'érythromycine. Ils ont ensuite été transformés avec l'ADN total de la souche 168 de B. subtilis dont le gène spoOA était interrompu par une "cassette Km<sup>R</sup>". Les clones transformants devenus résistants à la kanamycine et ayant toujours un phénotype Spo<sup>+</sup> ont été étudiés. L'un des clones portait un plasmide recombiné capable de compenser la mutation spoOA de B. subtilis. Ce plasmide était constitué du vecteur pHT304 et d'un fragment HindIII d'environ 2,4 kb (Fig. 11A).

Détermination de la séquence nucléotidique du gène spoOA de B. thuringiensis :

La séquence nucléotidique du fragment HindIII a été déterminée et a révélé la présence d'une phase ouverte de lecture de 804 pb susceptible de coder pour une protéine de 264 acides aminés homologue à la protéine SpoOA de B. subtilis. La séquence nucléotidique de 804 pb du gène spoOA de B. thuringiensis souche 407 est représentée Fig. 11B.

Interruption du gène spoOA de B. thuringiensis :

Un fragment d'ADN de 1,5 kb portant un gène aphIII, conférant la résistance à la kanamycine ("cassette Km<sup>R</sup>) a été inséré entre les deux sites HincII du gène spoOA (Fig. 11). Un fragment de 40 bp compris entre les positions 267 et 307 du gène spoOA a donc été remplacé par la "cassette Km<sup>R</sup>". Le fragment d'ADN HindIII d'environ 3,9 kb contenant le gène spoOA interrompu par la "cassette Km<sup>R</sup>" a été cloné dans le vecteur thermosensible pRN5101 (Villafane et al. 1987, J. Bacteriol. 169:4822-4829). Le plasmide résultant

(désigné pHT5120) a été introduit dans la souche de B.thuringiensis 407 Cry par électroporation. Le gène spoOA de la souche B.thuringiensis 407 Cry a été remplacé par la copie interrompue avec la "cassette Km<sup>R</sup>" par recombinaison génétique in vivo en utilisant le protocole précédemment décrit (Lereclus et al. 1992), Bio/Technology 10:418-421). La souche de B.thuringiensis résultante (désignée 407-OA::KM<sup>R</sup>) est résistante à la kanamycine (300 µg/ml) et ne produit aucune spore lorsqu'elle est cultivée en milieu HCT, habituellement favorable à la sporulation de B.thuringiensis. Une expérience d'hybridation ADN/ADN, réalisée avec le fragment HindIII de 2,4 kb comme sonde, révèle que le gène spoOA de la souche B.thuringiensis 407 Cry a bien été remplacé par la copie interrompue avec la "cassette Km<sup>R</sup>".

Production de la toxine CryIIIA dans la souche B.thuringiensis 407-OA::Km<sup>R</sup>:

Le plasmide pHT305P, portant le gène cryIIIA, a été introduit dans la souche B.thuringiensis 407-OA::Km<sup>R</sup> par électroporation. Le clone recombinant obtenu a été déposé à la CNCM le 03/05/1994 et pour lequel le numéro d'accession I-1412 lui a été assigné. Le clone recombinant obtenu a été cultivé à 30°C en milieu HCT + glucose 3 g/l ou en milieu LB (NaCl, 5 g/l ; extrait de levure, 5 g/l ; Bacto tryptone 10 g/l) pour estimer la production de toxines. Après environ 48 heures les bactéries contenaient un cristal visible par examen en microscopie optique. Ce cristal était de forme rhomboïdale, caractéristique des cristaux constitués par la protéine CryIIIA. Les cristaux produits par la souche B.thuringiensis 407-OA::Km<sup>R</sup> [pHT315] sont de taille importante et restent inclus dans les cellules plusieurs jours après que celles-ci aient cessé de se développer en milieu HCT ;

en milieu LB une partie des cellules lysent et les cristaux sont libérés. Les cristaux sont constitués de protéines d'environ 70 kDa (CryIIIA) spécifiquement toxiques pour les coléoptères.

REVENDICATIONS

1. Séquence d'ADN de contrôle de l'expression d'une séquence codante d'un gène dans une cellule hôte, ladite séquence étant caractérisée en ce qu'elle comprend un promoteur et une séquence nucléotidique ou région aval sous le contrôle dudit promoteur, et particulièrement située en aval dudit promoteur et en amont de ladite séquence codante, ladite séquence nucléotidique ou région aval comportant une région essentiellement complémentaire à l'extrémité 3' d'un ARN ribosomal bactérien, ladite séquence d'ADN étant susceptible d'intervenir pour augmenter l'expression de ladite séquence codante, placée en aval, dans un hôte cellulaire.

2. Séquence d'ADN selon la revendication 1, caractérisée par les propriétés suivantes :

- elle est comprise dans un enchaînement d'ADN d'une longueur de 1692 pb environ, délimité par les sites de restriction HindIII-PstI (fragment H<sub>2</sub>-P<sub>1</sub>), tel qu'obtenu par digestion partielle du fragment BamHI de 6 kb porté par le gène cryIIIA de Bacillus thuringiensis souche LM79 ;

- elle est susceptible d'intervenir dans le contrôle de l'expression d'une séquence codante de nucléotides placée en aval, dans un hôte cellulaire bactérien, particulièrement de type Bacillus thuringiensis et/ou Bacillus subtilis.

3. Séquence d'ADN selon la revendication 2 caractérisée en ce qu'elle correspond à l'enchaînement HindIII-PstI d'une longueur de 1692 pb environ (fragment H<sub>2</sub>-P<sub>1</sub>) tel qu'obtenu par digestion partielle du fragment BamHI de 6 kb portant le gène cryIIIA de Bacillus thuringiensis souche LM79.

4. Séquence d'ADN selon la revendication 1, caractérisée en ce qu'elle correspond à l'enchaînement de nucléotides suivant désigné par l'expression Seq N°1

compris entre les nucléotides 1 et 1692 de la séquence de la figure 3.

5. Séquence d'ADN selon l'une quelconque des revendications 2 à 4, caractérisée en ce qu'elle correspond à l'enchaînement délimité par les sites de restriction TagI-TagI.

6. Séquence d'ADN selon l'une quelconque des revendications 1 à 5, caractérisée ce qu'elle correspond à l'enchaînement de nucléotides suivant désigné par l'expression Seq N°2 compris entre les nucléotides 907 et 1559 de la séquence de la figure 3.

7. Séquence d'ADN selon l'une quelconque des revendications 2 à 4, caractérisée en ce qu'elle comprend le fragment TagI-PacI répondant à l'enchaînement Seq N°3, en association avec le fragment XmnI-TagI répondant à l'enchaînement Seq N°4.

8. Séquence d'ADN selon l'une quelconque des revendications 2 à 7, caractérisée en ce qu'il s'agit d'une séquence intervenant dans le contrôle de la transcription de la séquence codante de nucléotides ou d'une séquence intervenant pour augmenter l'expression d'une séquence codante.

9. Séquence d'ADN selon la revendication 1 à 18, caractérisée par :

- sa capacité d'hybridation dans des conditions non stringentes avec un enchaînement de nucléotides choisi parmi les enchaînements Seq N° 1, Seq N° 2 et Seq N° 3 et Seq N° 4 et par,

- sa capacité à intervenir dans le contrôle de l'expression, dans un hôte cellulaire, particulièrement du type Bacillus, d'une séquence codante de nucléotides placée en aval.

10. Séquence selon la revendication 1, caractérisée en ce que ladite séquence nucléotidique ou région aval est la séquence S2 isolée d'une bactérie, particulièrement d'une bactérie Gram<sup>+</sup>, plus particulièrement d'une bactérie de type Bacillus, ou

une séquence dérivée de ladite séquence S2, ladite séquence S2 étant caractérisée en ce qu'elle comporte une région essentiellement complémentaire à l'extrémité 3' de l'ARN 16S d'un ribosome bactérien, elle est susceptible d'intervenir au niveau post-transcriptionnel lors de l'expression de ladite séquence codante, notamment sans agir directement sur la traduction de la séquence codante à exprimer, particulièrement lorsque ladite séquence S2 est située entre ledit promoteur et ladite séquence codante d'un gène, afin d'augmenter l'expression de ladite séquence codante chez une bactérie, particulièrement une bactérie de type Bacillus.

11. Séquence d'ADN selon la revendication 10, caractérisée en ce que ladite séquence S2 comporte une première région suffisamment complémentaire à l'extrémité 3' de l'ARN 16S d'un ribosome bactérien, particulièrement d'une bactérie Gram<sup>+</sup>, plus particulièrement d'une bactérie de type Bacillus, pour que ladite première région soit susceptible de s'hybrider à un ribosome bactérien ou à une partie dudit ribosome, et une deuxième région en aval de la première région, ladite deuxième région comportant une séquence susceptible d'avoir un effet positif additionnel au niveau de l'expression de la séquence codante.

12. Séquence d'ADN selon la revendication 10, caractérisée en ce que la séquence nucléotidique ou région aval présente les propriétés suivantes :

- elle est contenue dans un enchaînement de nucléotides hybridant dans des conditions non stringentes avec le fragment d'ADN compris entre les nucléotides 1413 et 1559 de la séquence représentée à la figure 3 ;

- elle est susceptible d'intervenir dans le contrôle de l'expression de la séquence codante d'un gène dans une cellule hôte, notamment d'une séquence

codant pour un polypeptide de Bacillus toxique vis-à-vis des insectes ou d'une séquence codant pour un polypeptide exprimé pendant la phase végétative, pendant la phase stationnaire ou pendant la sporulation de Bacillus.

13. Séquence d'ADN selon l'une quelconque des revendications 2 à 12, caractérisée en ce que la séquence nucléotidique ou région aval correspond à l'enchaînement délimité par les sites de restriction XmnI-TaqI.

14. Séquence d'ADN selon la revendication 13, caractérisée en ce que la séquence nucléotidique ou région aval correspond à l'enchaînement Seq N° 4 compris entre les nucléotides 1179 et 1559 de la séquence de la figure 3, ou en ce qu'elle correspond à toute partie d'au moins 10 nucléotides naturellement consécutifs ou non de cet enchaînement, susceptible d'intervenir dans le contrôle de l'expression d'une séquence codante de nucléotides placée en aval dans un hôte cellulaire.

15. Séquence d'ADN selon la revendication 13, caractérisée en ce que la séquence nucléotidique ou région aval correspond à l'enchaînement Seq N° 5 correspondant au fragment comprenant les nucléotides 1179 à 1556 de la séquence représentée à la figure 3.

16. Séquence d'ADN selon la revendication 13, caractérisée en ce que la séquence nucléotidique ou région aval correspond à l'enchaînement Seq N° 11 correspondant au fragment comprenant les nucléotides 1413 à 1556 de la séquence représentée à la figure 3.

17. Séquence d'ADN selon la revendication 13, caractérisée en ce que la séquence nucléotidique ou région aval correspond à l'enchaînement Seq N° 8 correspondant au fragment comprenant les nucléotides 1413 à 1461 de la séquence représentée à la figure 3.

18. Séquence d'ADN selon la revendication 13, caractérisée en ce que la séquence nucléotidique ou

région aval correspond à l'enchaînement Seq N° 9 correspondant au fragment d'ADN suivant

5'-AGCTTGAAAGGAGGGATGCCTAAAAACGAAGAACTGCA-3'  
3'-ACTTCCCTCCCTACGGATTTGCTTCTTG-5'

19. Séquence selon la revendication 13, caractérisée en ce que la séquence nucléotidique ou région aval correspond à l'enchaînement Seq N° 10 correspondant au fragment d'ADN suivant :

5'-CTTGAAAGGAGGGATGCCTAAAAACGAAGAAC-3'  
3'-GAACTTCCCTCCCTACGGATTTGCTTCTTG-5'

20. Séquence d'ADN selon la revendication 13, caractérisée en ce que la séquence nucléotidique ou région aval correspond au fragment d'ADN suivant :

5'-GAAAGGAGG-3'  
3'-CTTCCCTCC-5'

21. Séquence d'ADN selon l'une quelconque des revendications 1 à 20, caractérisée en ce que le promoteur est inclus dans l'enchaînement délimité par les sites de restriction TagI-PacI.

22. Séquence d'ADN selon la revendication 21, caractérisée en ce que le promoteur correspond à l'enchaînement Seq N°3 compris entre les nucléotides 907 et 990 de la séquence de la figure 3, ou un variant comprenant les nucléotides 907 à 985, ou caractérisée en ce que le promoteur correspond à toute partie d'au moins 10 nucléotides naturellement consécutifs ou non de cet enchaînement, susceptible d'intervenir dans le contrôle de l'expression d'une séquence codante de nucléotides placée en aval dans un hôte cellulaire bactérien.

23. Séquence nucléotidique caractérisée en ce qu'elle comporte une région essentiellement complémentaire à l'extrémité 3' de l'ARN 16S d'un

ribosome bactérien, ladite séquence étant susceptible d'intervenir au niveau post-transcriptionnel lors de l'expression de la séquence codante d'un gène, notamment sans agir directement sur la traduction de la séquence codante à exprimer, ladite séquence nucléotidique étant sous le contrôle d'un promoteur contrôlant ladite séquence codante.

24. Séquence nucléotidique selon la revendication 23, caractérisée en ce que ladite séquence est isolée d'une bactérie, particulièrement d'une bactérie Gram<sup>+</sup> et plus particulièrement d'une bactérie de type Bacillus.

25. Séquence nucléotidique selon la revendication 23, caractérisée en ce que ladite séquence est une séquence S2 comportant une première région suffisamment complémentaire à l'extrémité 3' de l'ARN 16S d'un ribosome bactérien, particulièrement d'une bactérie Gram<sup>+</sup>, plus particulièrement d'une bactérie de type Bacillus pour que ladite première région soit susceptible de s'hybrider à un ribosome bactérien ou une partie d'un ribosome bactérien et une deuxième région en aval de la première région, ladite deuxième région comportant une séquence susceptible d'avoir un effet positif additionnel au niveau de l'expression de la séquence codante.

26. Séquence nucléotidique selon la revendication 24, caractérisée en ce que la séquence nucléotidique correspond à l'enchaînement Seq N° 5 correspondant au fragment comprenant les nucléotides 1179 à 1556 de la séquence représentée à la figure 3.

27. Séquence nucléotidique selon la revendication 24, caractérisée en ce que la séquence nucléotidique correspond à l'enchaînement Seq N° 11 correspondant au fragment comprenant les nucléotides 1413 à 1556 de la séquence représentée à la figure 3.

28. Séquence nucléotidique selon la revendication 24, caractérisée en ce que la séquence

nucléotidique correspond à l'enchaînement Seq N° 8 correspondant au fragment comprenant les nucléotides 1413 à 1461 de la séquence représentée à la figure 3.

29. Séquence nucléotidique selon la revendication 24, caractérisée en ce que la séquence nucléotidique ou région aval correspond à l'enchaînement Seq N° 9 correspondant au fragment d'ADN suivant

5'-AGCTTGAAAGGAGGGATGCCTAAAAACGAAGAACTGCA-3'  
3'-ACTTCCCTCCCTACGGATTGGCTTCTTG-5'

30. Séquence nucléotidique selon la revendication 24, caractérisée en ce que la séquence nucléotidique ou région aval correspond à l'enchaînement Seq N° 10 correspondant au fragment d'ADN suivant

5'-CTTGAAAGGAGGGATGCCTAAAAACGAAGAAC-3'  
3'-GAACTTTCCCTCCCTACGGATTGGCTTCTTG-5'

31. Séquence nucléotidique selon la revendication 24, caractérisée en ce que la séquence nucléotidique correspond au fragment d'ADN suivant :

5'-GAAAGGAGG-3'  
3'-CTTCCTCC-5'

32. Séquence d'ADN recombinant, caractérisée en ce qu'elle comprend une séquence codante déterminée sous le contrôle d'une séquence d'ADN selon l'une quelconque des revendications 1 à 22.

33. Vecteur d'expression caractérisé en ce qu'il est modifié en l'un de ses sites, par une séquence d'ADN selon l'une quelconque des revendications 1 à 22, de façon telle que ladite séquence d'ADN intervient dans le contrôle de l'expression d'une séquence codante de nucléotides déterminée.

34. Vecteur d'expression selon la revendication 32, caractérisé en ce qu'il s'agit d'un plasmide, par exemple d'un plasmide intégratif, ou d'un plasmide répliquatif.

35. Vecteur d'expression selon la revendication 34, caractérisé en ce qu'il s'agit du plasmide pHT7902'lacZ déposé à la CNCM le 20 avril 1993 sous le n° I-1301.

36. Séquence d'ADN recombinant selon la revendication 32 ou vecteur d'expression selon la revendication 33 ou 34, caractérisée en ce que la séquence codante de nucléotides qu'il contient est une séquence毒ique vis-à-vis des insectes et en particulier une séquence à activité larvicide, par exemple la séquence codante du gène cryIIIA de B.thuringiensis.

37. Séquence d'ADN recombinant selon la revendication 32 ou vecteur d'expression selon la revendication 33 ou 34, caractérisée en ce que la séquence codante de nucléotides qu'il contient est une séquence codant pour une enzyme.

38. Séquence d'ADN recombinant selon la revendication 32 ou vecteur d'expression selon la revendication 33 ou 34, caractérisée en ce que la séquence codante de nucléotides qu'il contient est une séquence codant pour un antigène.

39. Hôte cellulaire recombinant, caractérisé en ce qu'il est modifié par une séquence d'ADN selon l'une quelconque des revendications 1 à 32 ou par un vecteur d'expression selon l'une quelconque des revendications 33 ou 34.

40. Hôte cellulaire selon la revendication 39, caractérisé en ce qu'il s'agit d'un Bacillus, par exemple de B. thuringiensis ou de B. subtilis.

41. Hôte cellulaire selon la revendication 40, caractérisé en ce qu'il s'agit d'une souche asporogène de Bacillus, par exemple d'une souche de B.

subtilis asporogène, en particulier une souche capable d'exprimer la séquence codante de la séquence d'ADN selon l'une quelconque des revendications 14 à 18, pendant la phase stationnaire de croissance.

42. Hôte cellulaire, caractérisé en ce qu'il s'agit de la souche 407-OA::Km<sup>R</sup> (pHT305D) déposée à la CNCM le 3 mai 1994 sous le n° I-1412.

43 Procédé de production d'une protéine recombinante codée par une séquence de nucléotides déterminée, ledit procédé étant caractérisé en ce qu'il comprend :

- l'introduction dans un hôte cellulaire d'un vecteur selon l'une quelconque des revendications 33 à 38,

- la croissance dudit hôte cellulaire à des conditions permettant l'expression de ladite séquence de nucléotides déterminée, et

- la récupération de la protéine recombinante.

1/13

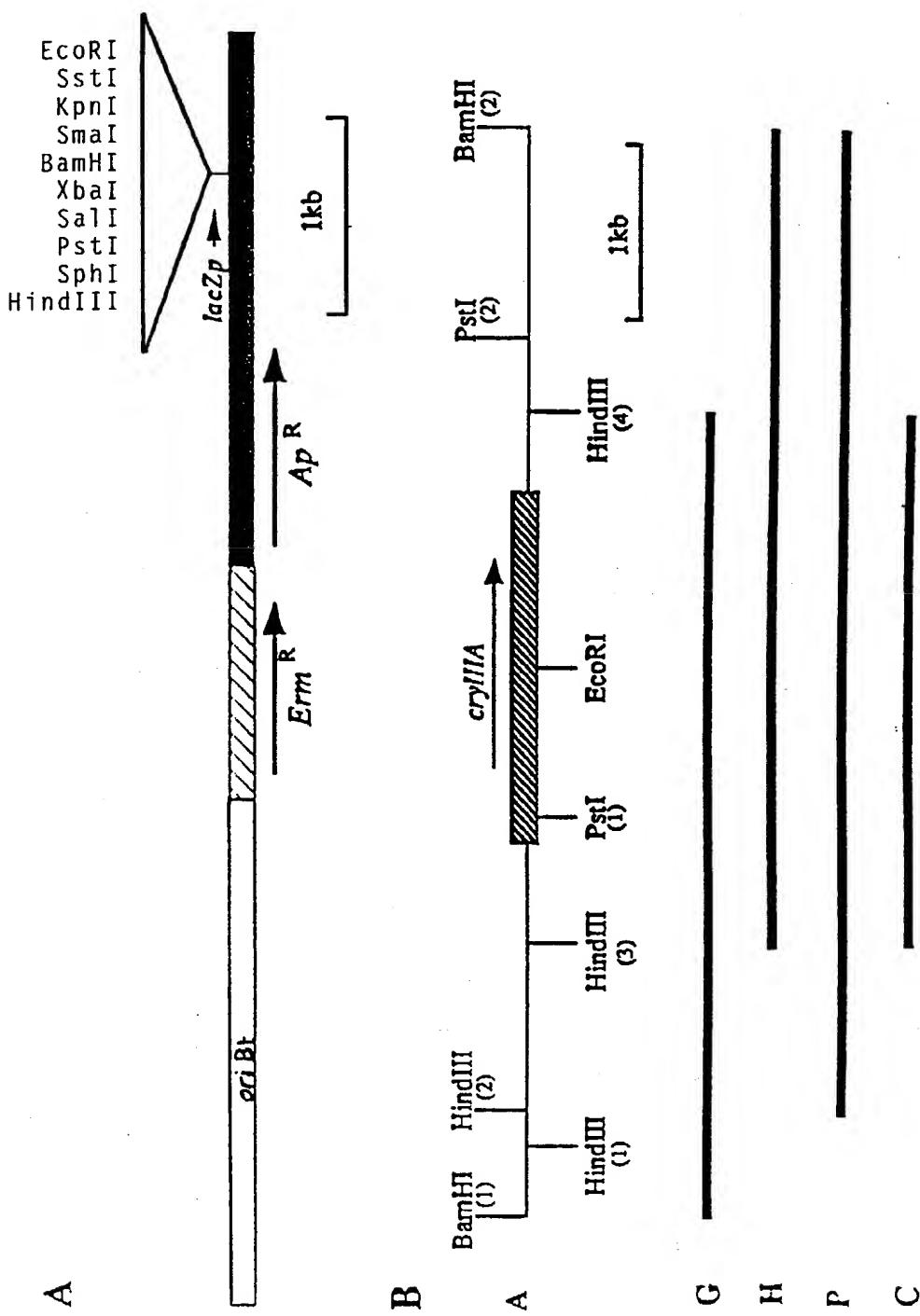


FIGURE I

2/13

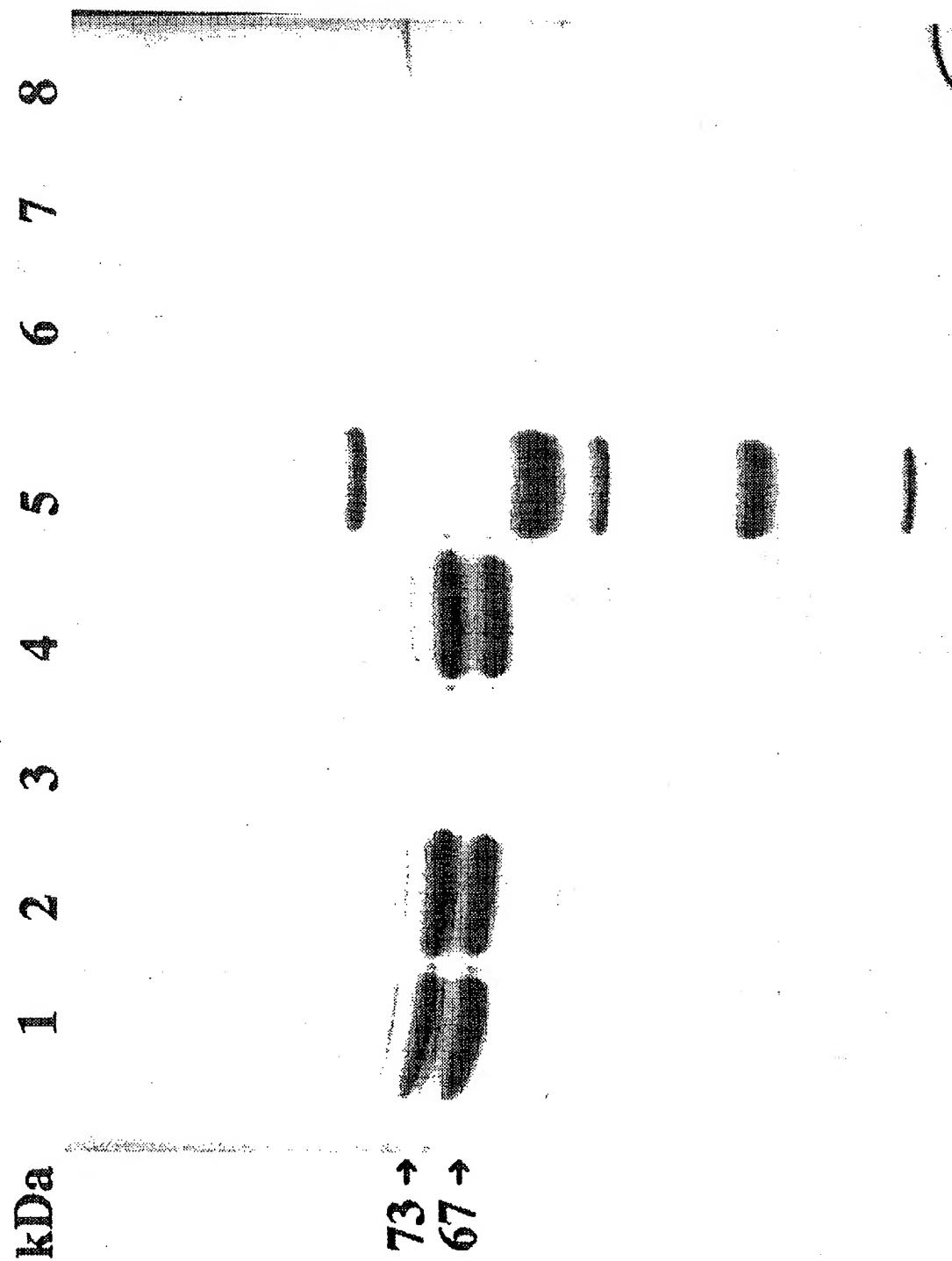


FIGURE 2

FEUILLE DE REMPLACEMENT (REGLE 26)

3/13

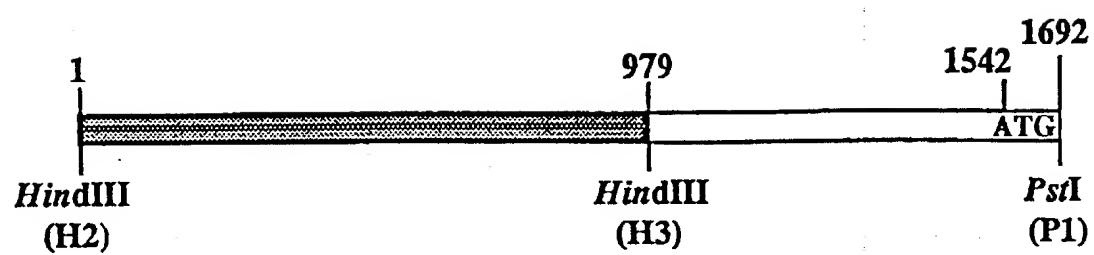


FIGURE 3A

4/13

1 AAGCTTTCAG TGAAGTAGT GATTATAACGG AGATGAAAT TCGTACACTG TTACCGAGAA  
 61 GGAAACGCCG ACGAAAGCCGT AGCATCGGT GAGTAACGAA TATCTCTAACG  
 121 GTGTACTGGG GCTTTACTGA GACTAGAAAG GAGAGTTTC  
 181 GATAAAAGTG TCAGCCATT GATAAGTCTC ATTCTCAATAA CCTCTTCACCTT GAAAAGTGCA  
 241 GGAAGCTGCT TGAGAGGGAA AACCTCACGA ACAGTTCTTA TGGGGAGAGA CTGGAAACAG  
 301 GTCACAAATTG ATACCTCGCT AATCTTAA CCGACAAAGT TTTTAAAC CGTGGAAAGTC  
 361 ATAATAACCT GGATATTGTG ATTITATAAA AGTTAACAAA TGGTTTATAT TAAGACAGTC  
 421 ATAACCAAA GATTTTCTT CAAAGCTAC GATAAGAAAA ATTTCACTAG AAATTAGTTA  
 481 TACAAGCATT TTGTAAGAAT TATAAAAG ATAAATCCTG CTATTACGAG ATTAGTAGGA  
 541 TGATATTTG AAAATTTTT TATCTATTCG ATTAAATAA TTATGAAATT TTACATAAAC  
 601 CTCATAAGAA AAAATACTAT CTATACTATT TTAAAGAATT TATTAGAATA AGCGGATTCA  
 661 AAATAGCCCT GCCCATAAAA GTACCTAGC AGTAGAAAGT TTGACCAAAA TTAAAAAAAT  
 721 ACCCAATCAA GAGAATATTTC TTAATTACAA TAGTTTTCG GAGGAACATA TTGATTGAAA  
 781 TTTAATAAAAT TTAGTCCCAA ATTAAAGA AATTAAAGT TTTCATATT TTATGAACTA  
 841 ACAAGAAATAA AAATTGTGTT TATTATATT TCTTGTAAAT TATTGATAA AGAGATATAT  
 901 TTTGGCGA AACGTAAGAT GAAACCTTAG ATAAAGTGC TTTTTTGTG GCAATTGAG  
 961 ATTATTAAAT GTTAAGCTTA ATTAAAGATA ATATCTTGA ATTGTAAACGC CCCTCAAAAG  
 1021 TAAGAACTAC AAAAAGAA TAGCTTATAT AGAAATATGT TTGAACCTTC TTCAAGATTAC  
 1081 AAATATATTG GGACGGACTC TACCTCAAAT GCTTATCTAA CTATAGAATG ACATACAAGC  
 1141 ACAACCTTGA AAATTGAAA ATATAACTAC CAATGAACCTT GTTCATGTGA ATTATCGCTG  
 1201 TATTTAATT TCTCAATTCA ATATAATAA TGCCAATACA TTGTTACAG TAGAAATTAA  
 1261 GACACCTTG ATAGCCTTAC TATACTAAC ATGATGTAGT ATAAATGAA TATGTAATAA  
 1321 TATTTATGAT AGAAGGGAC TTATTATAA TCATTACATA TTTTCTATT GGAATGATTA  
 1381 AGATTCAAT AGAATAGTGT ATAATTATT TATCTGAAA GGAGGGATGC CTAAAACGA  
 1441 AGAACATTAA AAACATATAT TTGCACCGTC TAATGGATT ATGAAATACT ATTTATCAG  
 1501 TTTGAAAATT ATGTATATG ATAAGAAAGG GAGGAAGAAA AATGAAATCCG AACAAATCGA  
 1560 AGTGAACAT GATACAATA AAAACTACT GAAAATATT GAGGTGCCA ACTAACCAT  
 1614 GTTCAATAT CTTTAGCG GAAACTCCA AATCCAAACA CTAGAAGAT TTAAATTAT  
 1668 AAAGAGTT TTAAGAATG ACTGCAG

FIGURE 3B

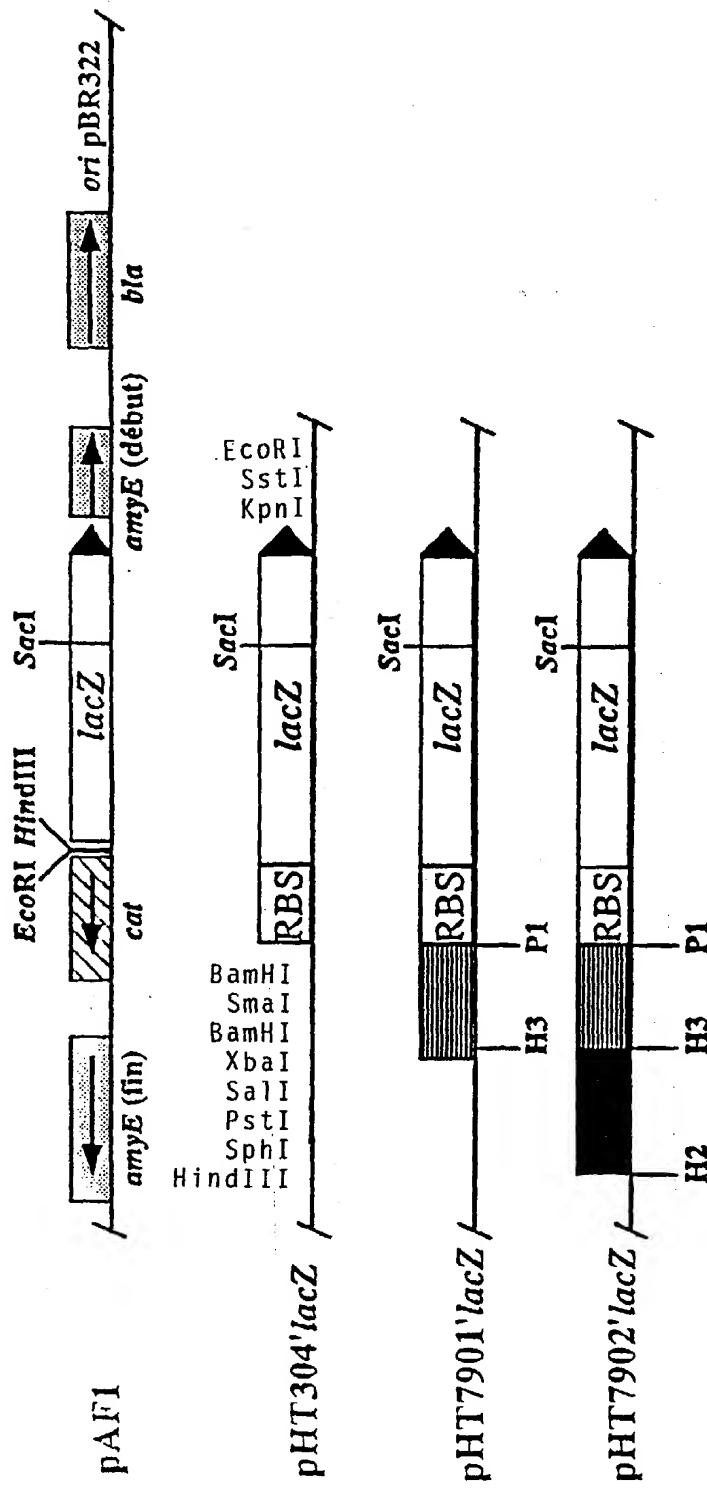


FIGURE 4

6/13

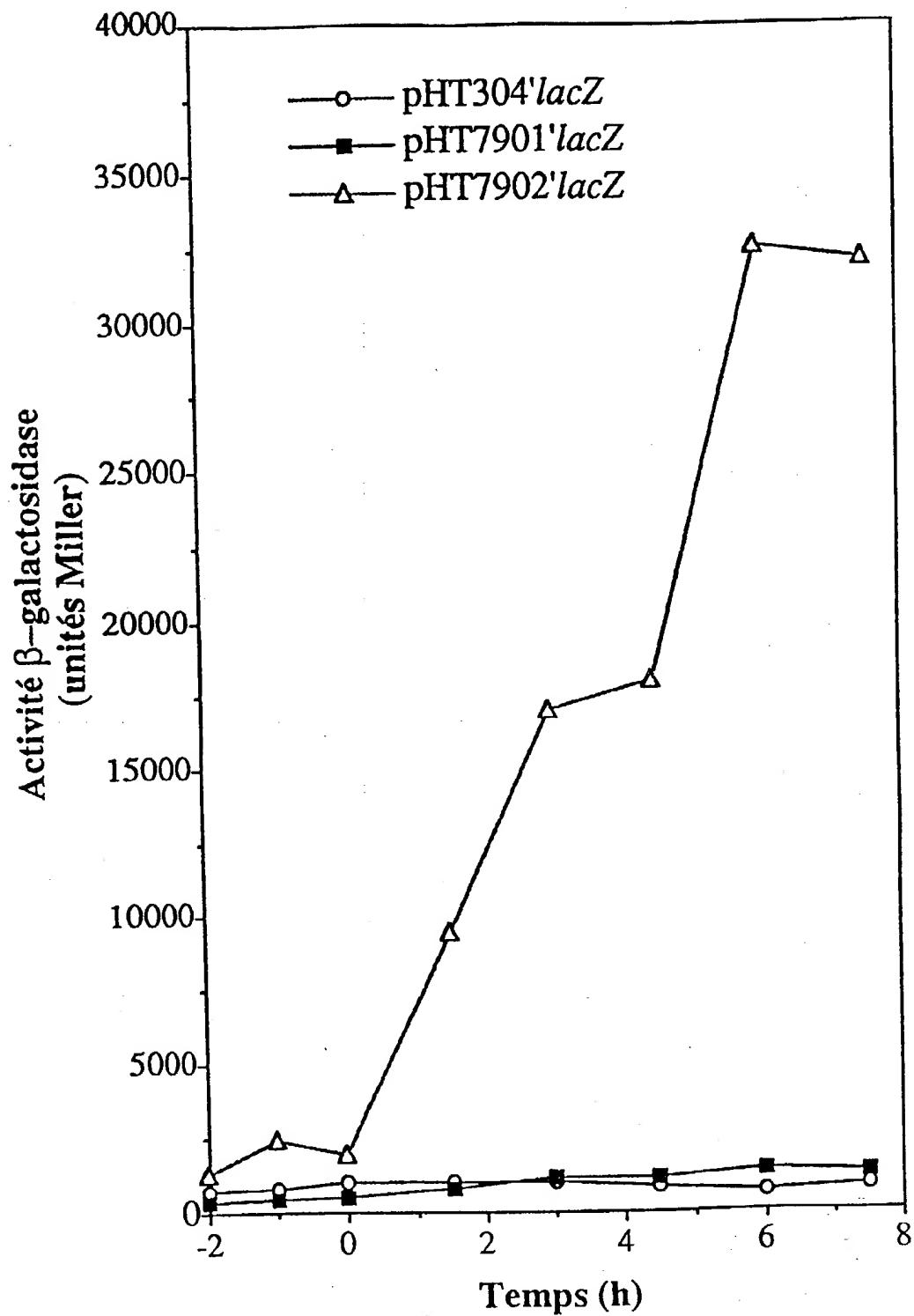


FIGURE 5

7/13

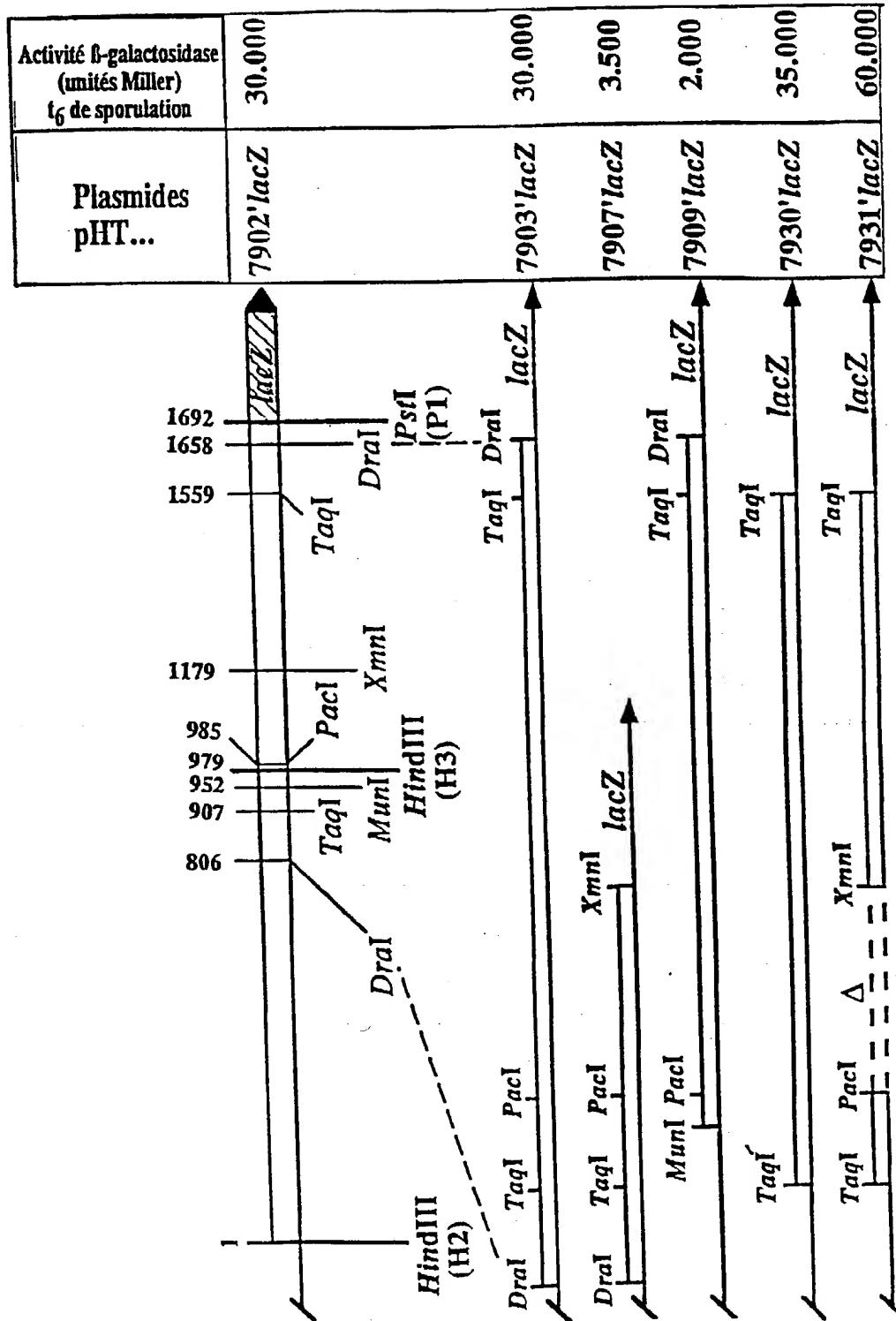


FIGURE 6

8/13

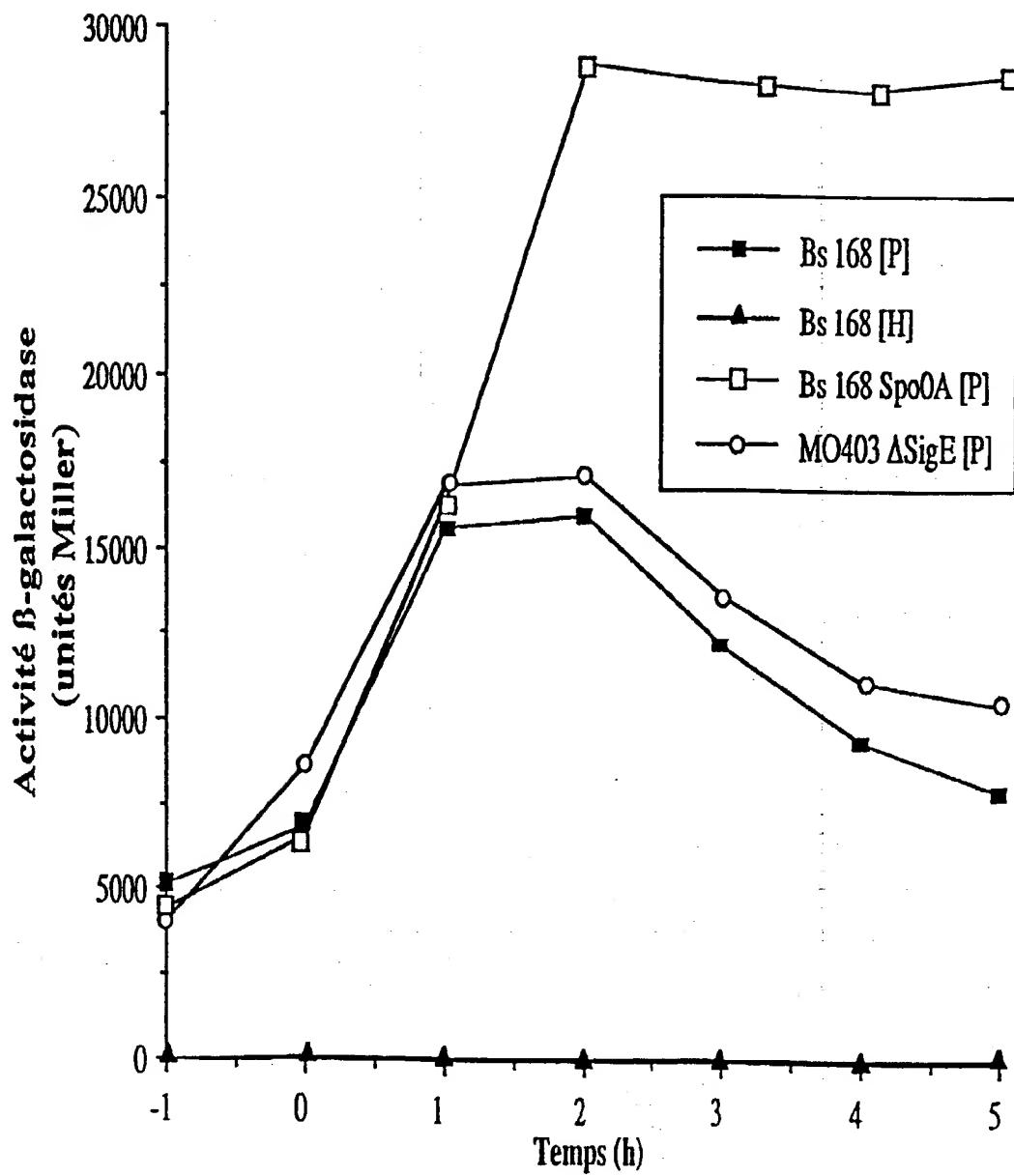


FIGURE 7

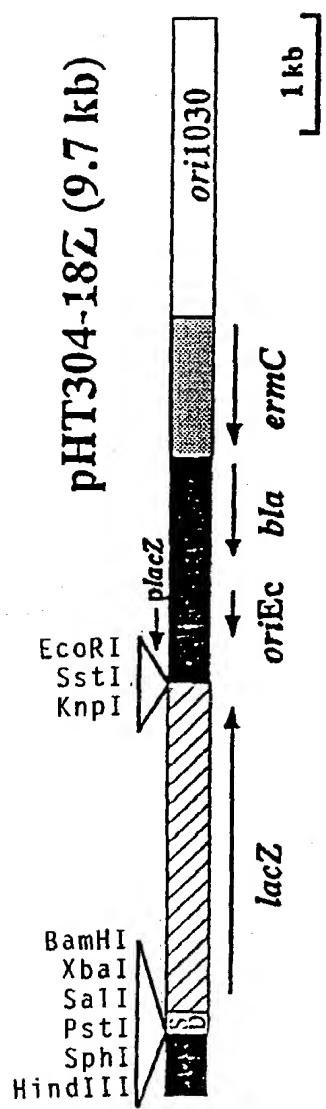


Fig. 8A

10/13

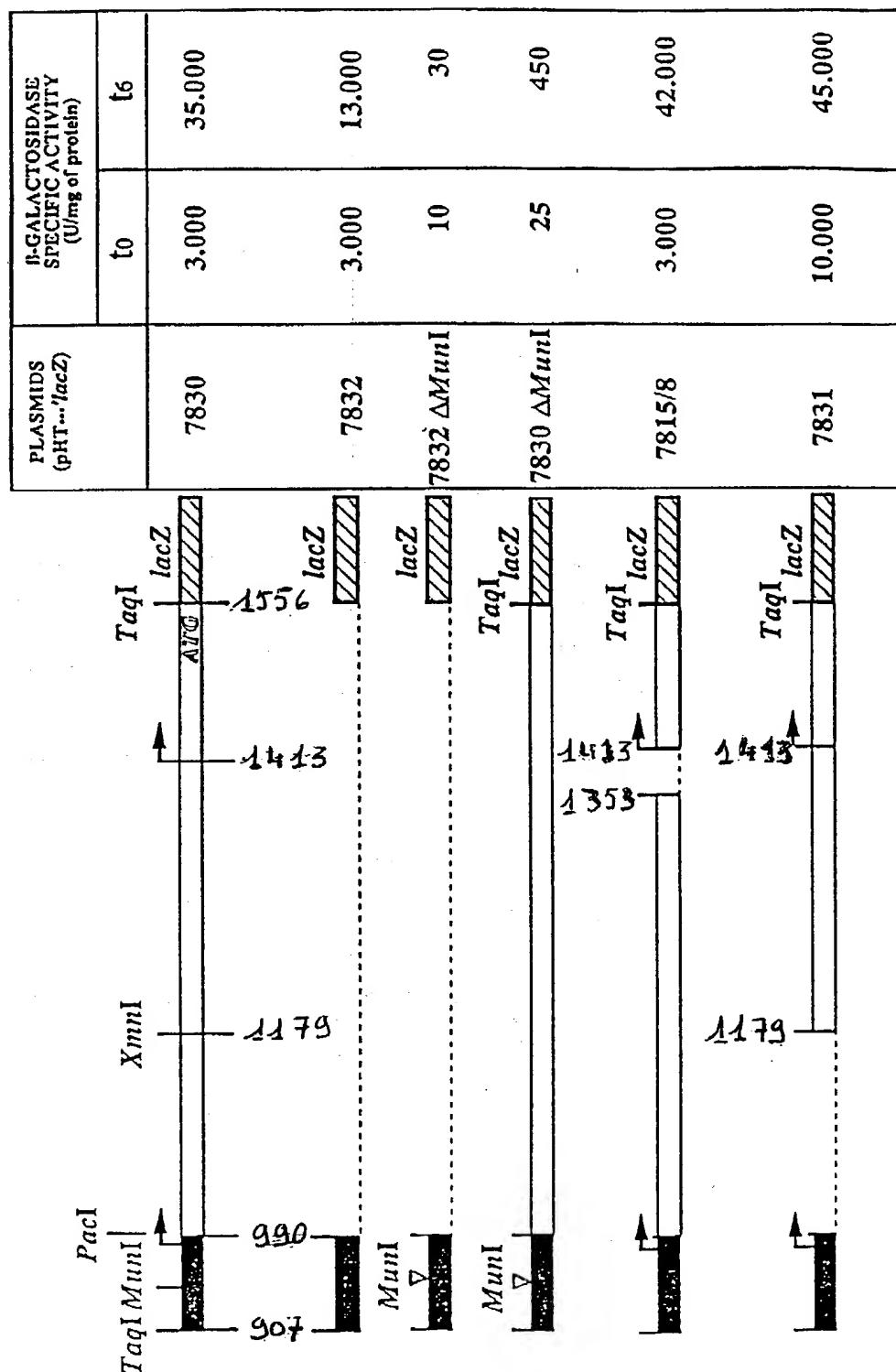


Fig. 8B

11/13

A C G T - 1

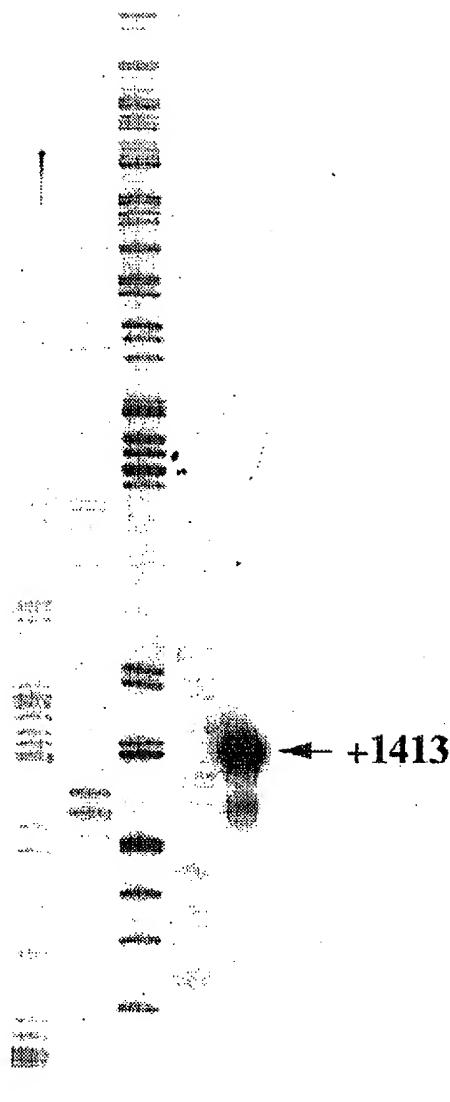
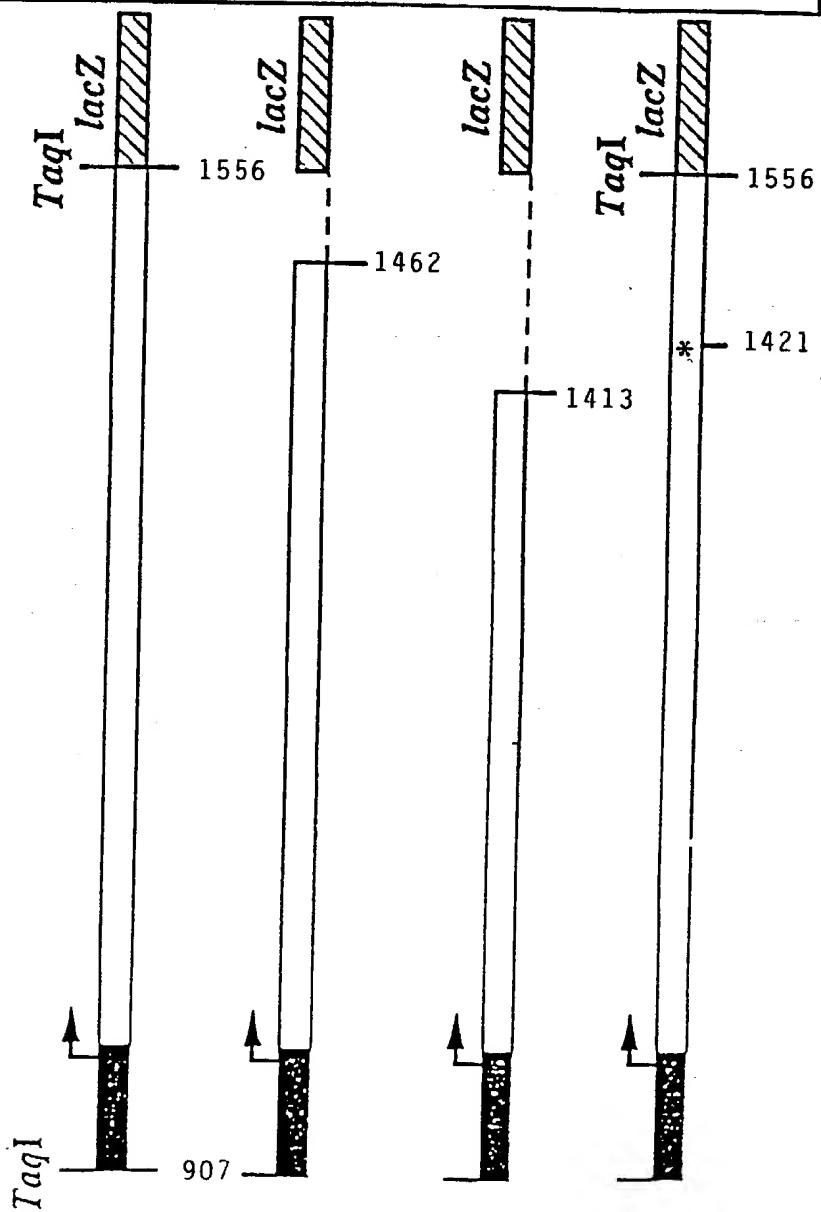


FIGURE 9

12/13

| PLASMIDES<br>(pHT...' <i>lacZ</i> ) | $\beta$ -GALACTOSIDASE<br>ACTIVITÉ SPÉCIFIQUE<br>(U/mg de protéines) | t <sub>3</sub> |
|-------------------------------------|--|----------------|
| 7830                                | 50.000   |                |
| 7816                                | 25.000   |                |
| 7805                                | 5.000  |                |
| 7830Rm                              | 5.000  |                |



13/13

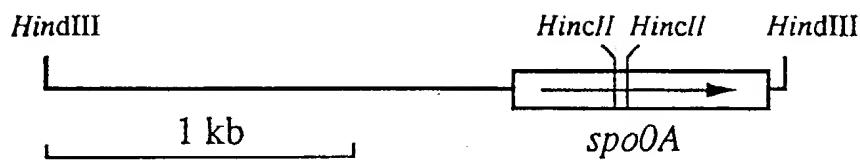


Fig. 11A

|     | 10<br>*  | 20<br>* | 30<br>* | 40<br>* |
|-----|--|---------|---------|---------|
| 1   | GGAGGAAAAGCTGTGGAGAAAATTAAAGTATGTCTTGTGG         |         |         |         |
| 41  | ATGATAATAAAGAATTAGTATCAATGTTAGAGAGCTATGT         |         |         |         |
| 81  | AGCCGCCCAAGATGATATGGAAGTAATCGGTACTGCTTAT         |         |         |         |
| 121 | AATGGTCAAGAGTGTAACTTATTAAACAGATAAGCAAC           |         |         |         |
| 161 | CTGATGTACTCGTTTAGACATTATTATGCCACACTTAGA          |         |         |         |
| 201 | TGGTTTAGCTGTATTGGAAAAAAATGCGACATATTGAAAGG        |         |         |         |
| 241 | TTAAAACAGCCTAGCGTAATTAT <u>GTTGACAGCATCGGGC</u>  |         |         |         |
| 281 | AAGAAGATGTGACGAAAAAAGCAG <u>TTGACTTAGGTGCCTC</u> |         |         |         |
| 321 | GTATTTCATATTAAAACCATTGATATGGAGAATTAAACG          |         |         |         |
| 361 | AGTCATATTCGTCAAGTGAGTGGTAAAGCAAACGCTATGA         |         |         |         |
| 401 | TTAACCGTCCACTACCACATTCCGATCAGCAACAAACAGT         |         |         |         |
| 441 | AGATGGAAAACC <u>GAAAAACTTAGATGCGAGTATTACGAGT</u> |         |         |         |
| 481 | ATCATTCA <u>TGAAATTGGTGTACCCGCTCATATTAAAGGAT</u> |         |         |         |
| 521 | ATATGTATTACGAGAAGCAATCTCCATGGTATACAATGA          |         |         |         |
| 561 | TATCGAATTATTAGGATCGATTACGAAAGTATTGTATCCA         |         |         |         |
| 601 | GATATCGCAAAGAAATATAATACAAACAGCCAGCCGTGTGG        |         |         |         |
| 641 | AGCGCGCAATT <u>CGTCACGCAATTGAAGTAGCTTGGAGCCG</u> |         |         |         |
| 681 | TGGGAATATTGATTCTATT <u>CGTCCTTATTGCGTTATACA</u>  |         |         |         |
| 721 | GTATCCATGTCAAAAGCAAAACCTACGA <u>ACTCTGAGTTA</u>  |         |         |         |
| 761 | TCGCAATGGTTGCGGATAAGCTGAGACTTGAACATAAAGC         |         |         |         |
| 801 | TAGT...  |         |         |         |

Fig. 11B